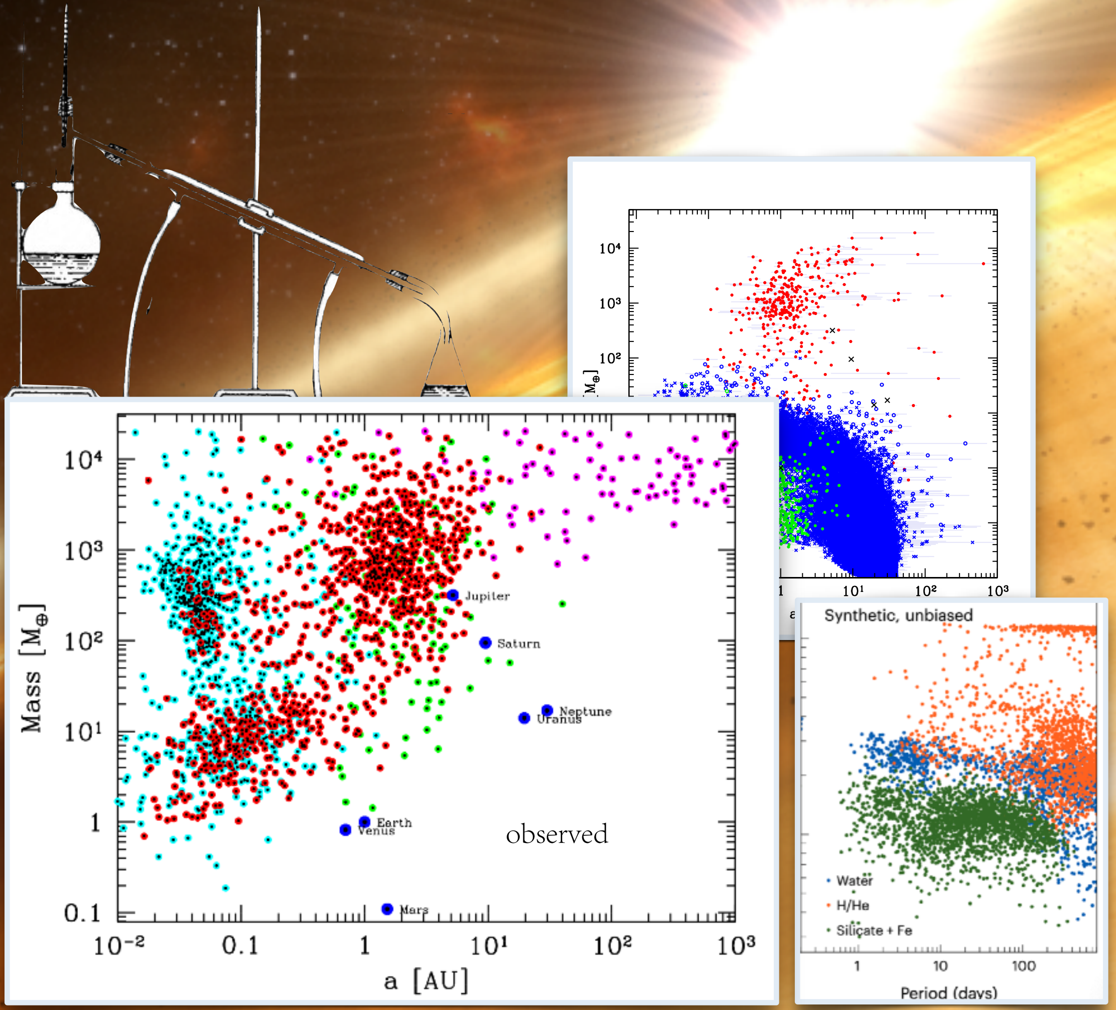
Synthesis of Exoplanet Demographics - an overview

2025 Sagan Summer Workshop

NASA Exoplanet Science Institute

California Institute of Technology

Pasadena 21.07.2025

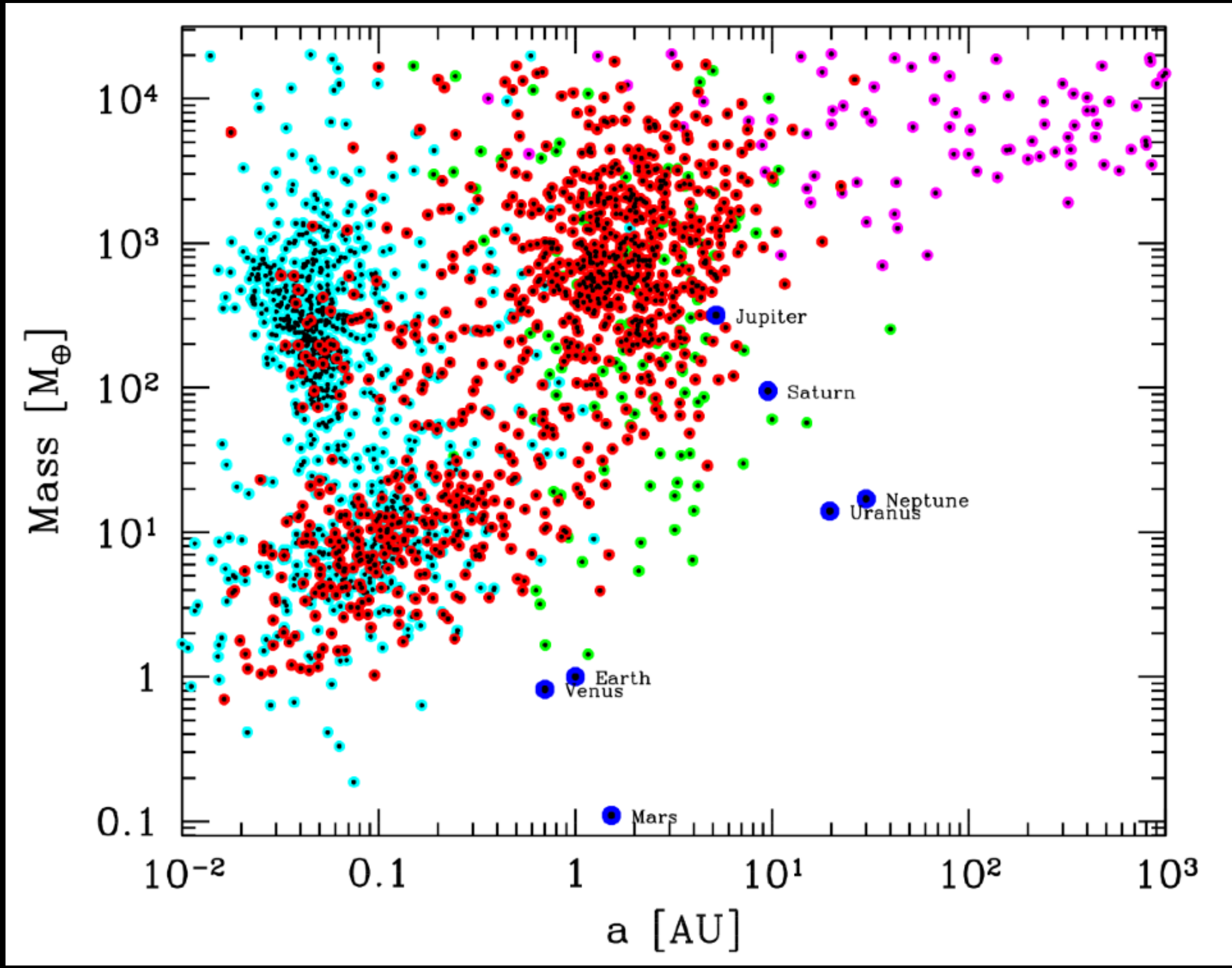


Christoph Mordasini
Division of Space Research and Planetary Sciences, Physics Institute, University of Bern, Switzerland

UNIVERSITÄT BERN
PlanetS
Swiss National Science Foundation

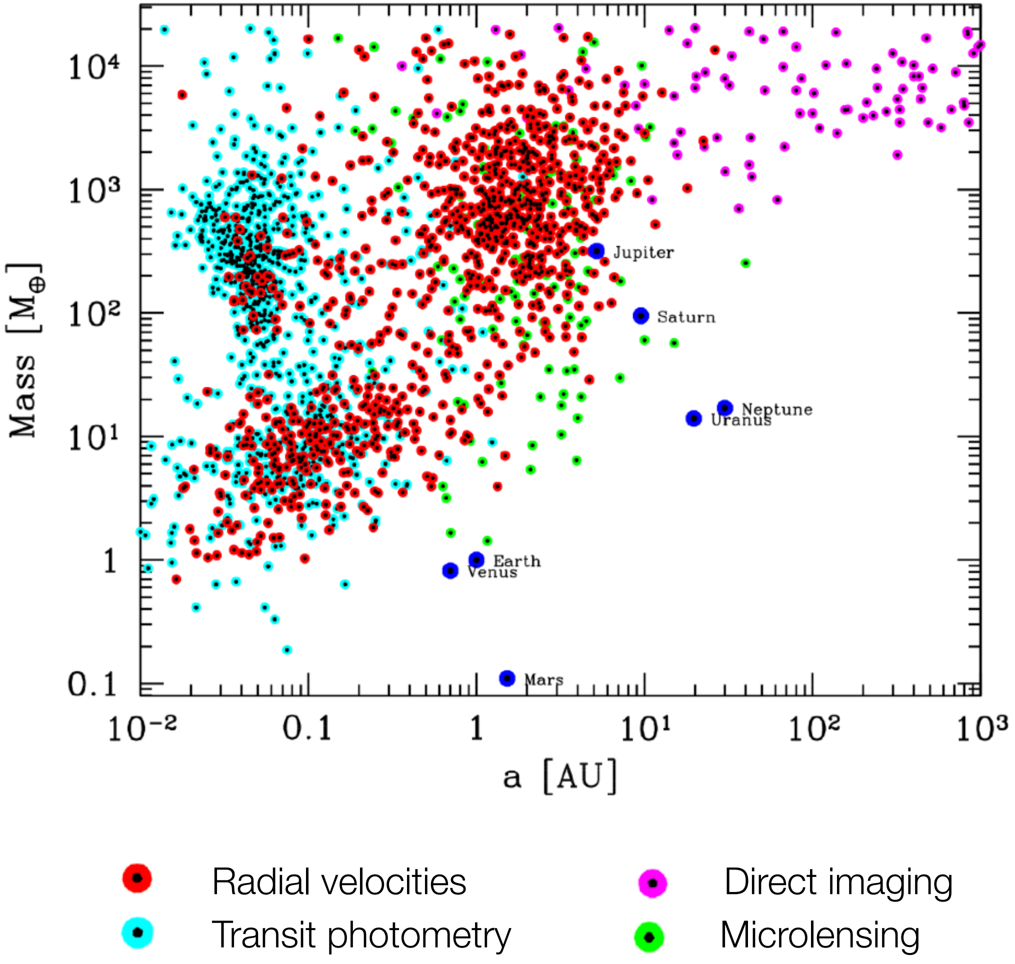
Talk overview

1. Observational motivation
2. Planetary populations synthesis method
3. Global models
4. Comparison with observations
5. Conclusions and future steps



1.
Observational motivation

Observational motivation



30 years of exoplanet studies

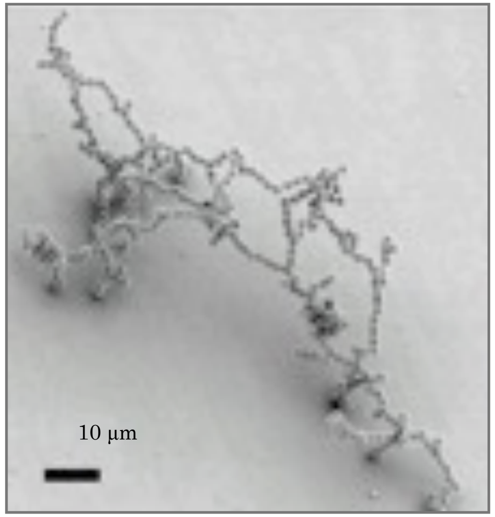
- Enormous increase in observational data on exoplanets since 1995. Detections from ground and space (HARPS, Kepler, NGTS, WASP, SPHERE, GPI, CARMENES, TESS, CHEOPS, ESPRESSO, NIRPS, JWST, …)

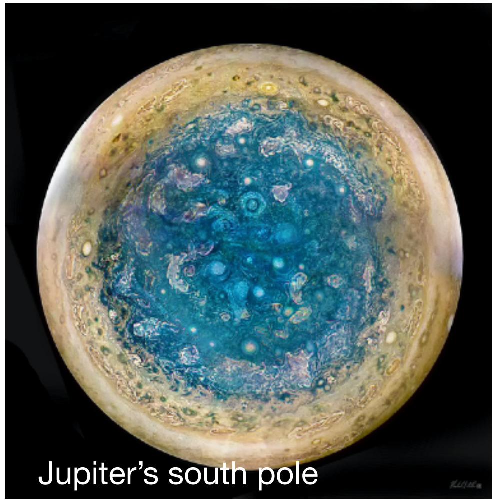
- More to come soon (Gaia DR4, PLATO, Roman ST, ARIEL, ELT, …)

Diversity in exoplanet properties

We would like to use all these demographics to better understand planet formation and evolution. But the field remains observationally driven, theory struggles to keep up. Why?

Challenges in planet formation and evolution theory





Planet formation is a complex process

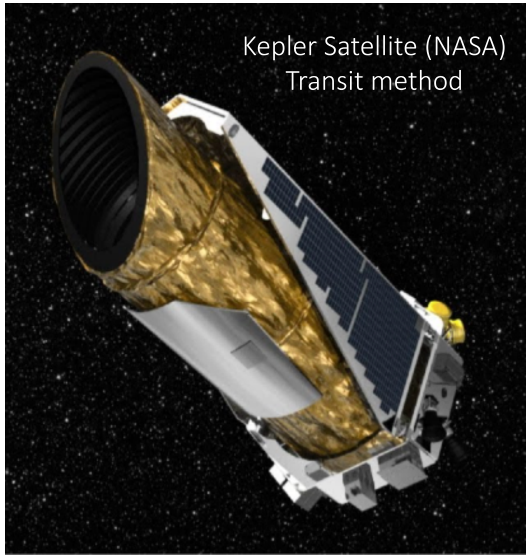
- Huge range in spatial scales: dust grains to giant planets
- Millions of dynamical timescales
- Multiple input physics: gravity, hydrodynamics, thermodynamics, radiative transport, magnetic fields, high-pressure physics,…
- Strong non-linear mechanisms and feedbacks

- Laboratory experiments only for special regimes
- Complete 3D radiation-magnetohydrodyamic numerical simulations too expensive

Cannot build theory based on first principles of physic only.

⇒ Theory needs observational guidance via comparison of observations and theoretical predictions

Comparing theory and observations



Compelling comparisons not so easy in practice:

- theoretical models for specific processes: difficult to test directly with observations. Each physical mechanism intermingles with many others. Only result of non-linearly combined action of all mechanisms observable.
- often only limited knowledge about an individual exoplanet system (like period and radius / minimum mass).

But: very high number of exoplanets: they can be treated as a population.

- demographical constraints
- data from many different techniques probing different parameters spaces: stringent constraints on theoretical models by combining M, a, e, R, L, spectra, …

We need a tool to use this wealth of constraints.

Population synthesis as a tool

Population synthesis is a tool to:

- use the exoplanets demographics to constrain planet formation and evolution models
- test the observational implications of theoretical concepts
- predict the yield of future instrumentation and surveys
- provide a link between theory and observations

Statistical approach rather than comparing individual systems

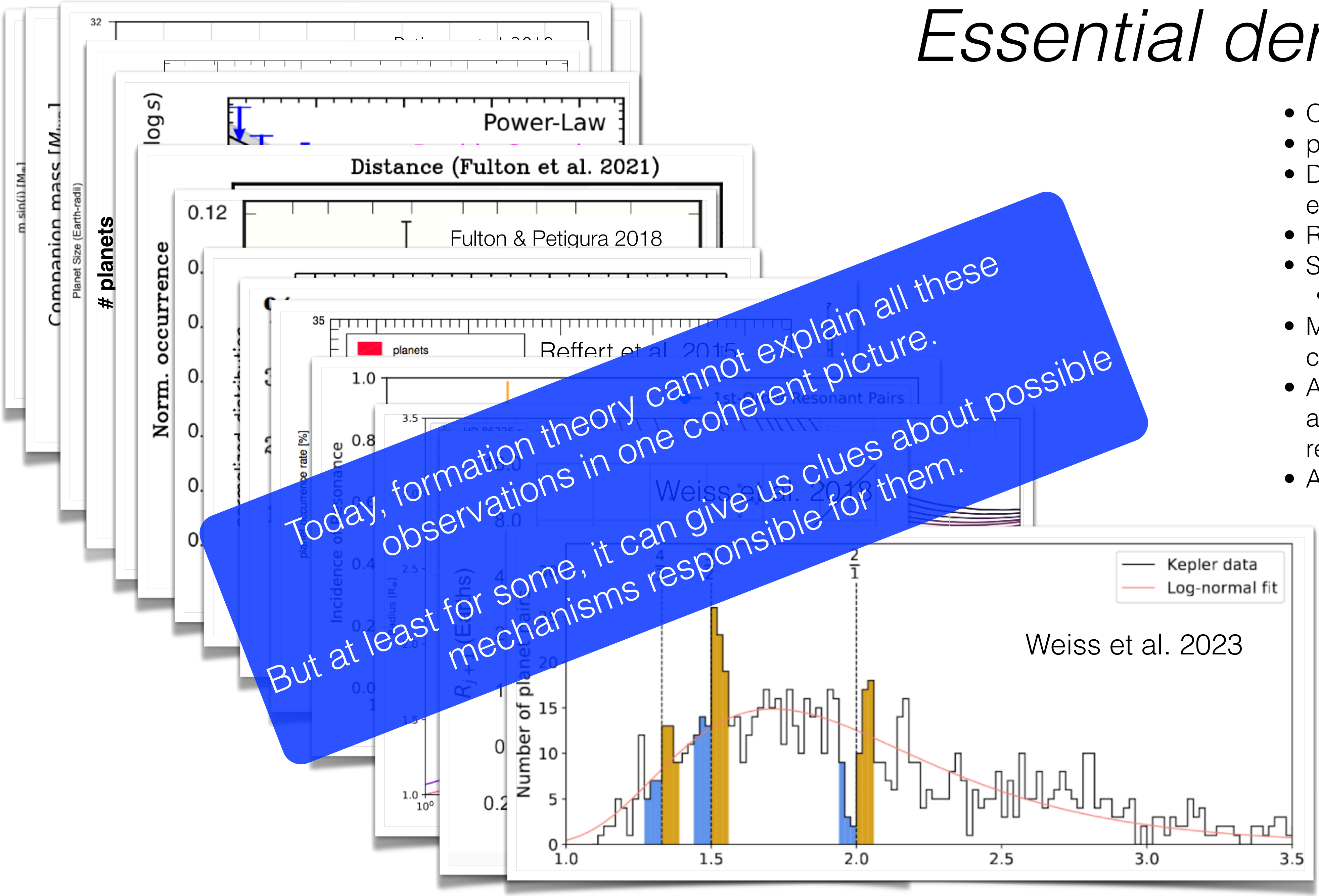
- need to compute the formation of many planetary systems
- the approach and the physics must be simplified (typically low-dimensional)
- but it must capture the key effects

⇒ builds on all detailed studies of specific physical mechanisms, combining them into a global end-to-end formation & evolution model

- depends on / reflects the general progress of the field

Essential demographics

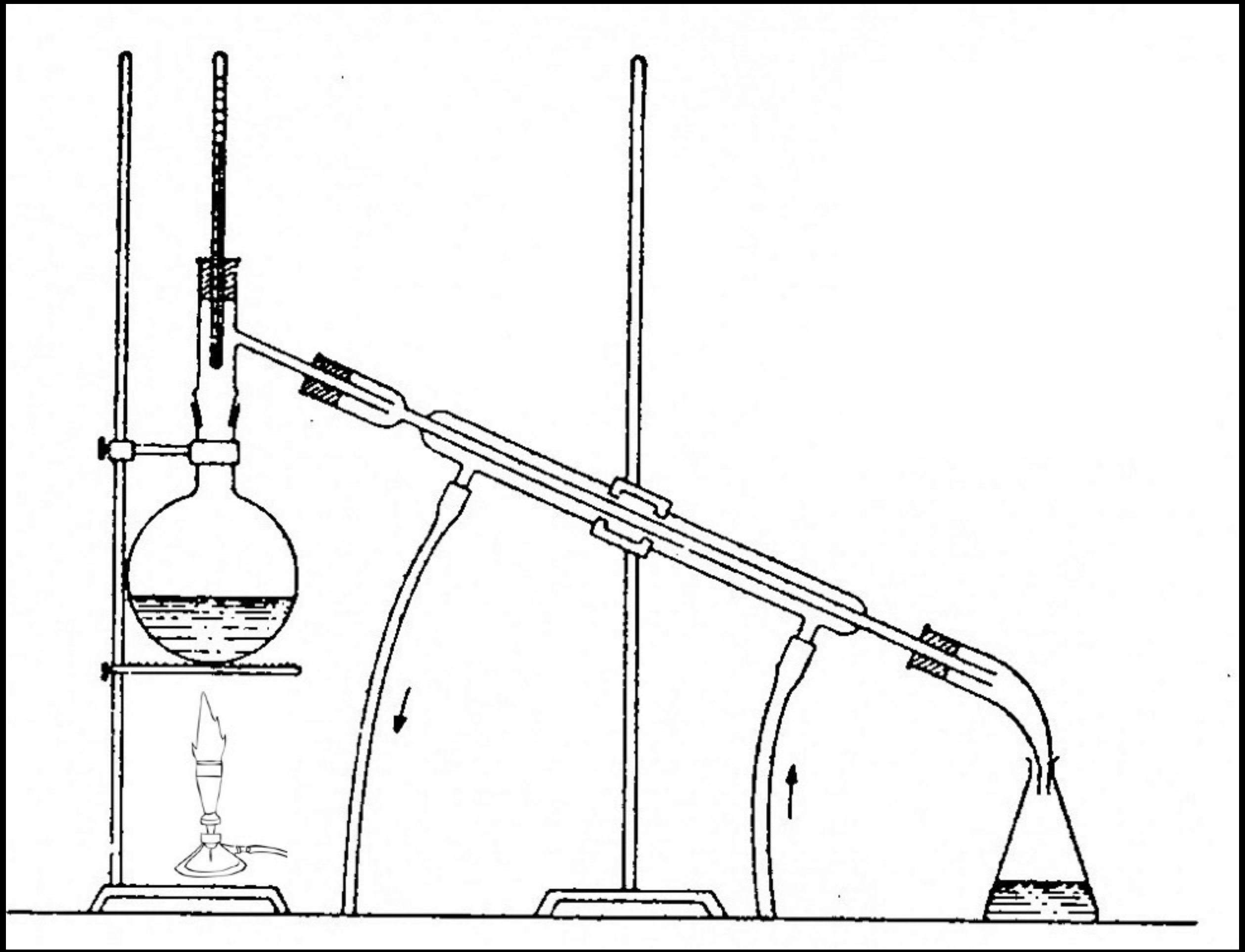
- Occurrence rates of planet types
- period-M/R/mag diagrams
- Distributions: mass, distance, eccentricity
- Radius distribution, radius valley
- Stellar dependencies
  - [Fe/H], stellar mass, age
- Mass-radius relation, bulk composition
- Architecture (multiplicity, peas-in-a-pod, resonances, SE-CJ relation)
- Atmospheric composition



Today, formation theory cannot explain all these observations in one coherent picture.

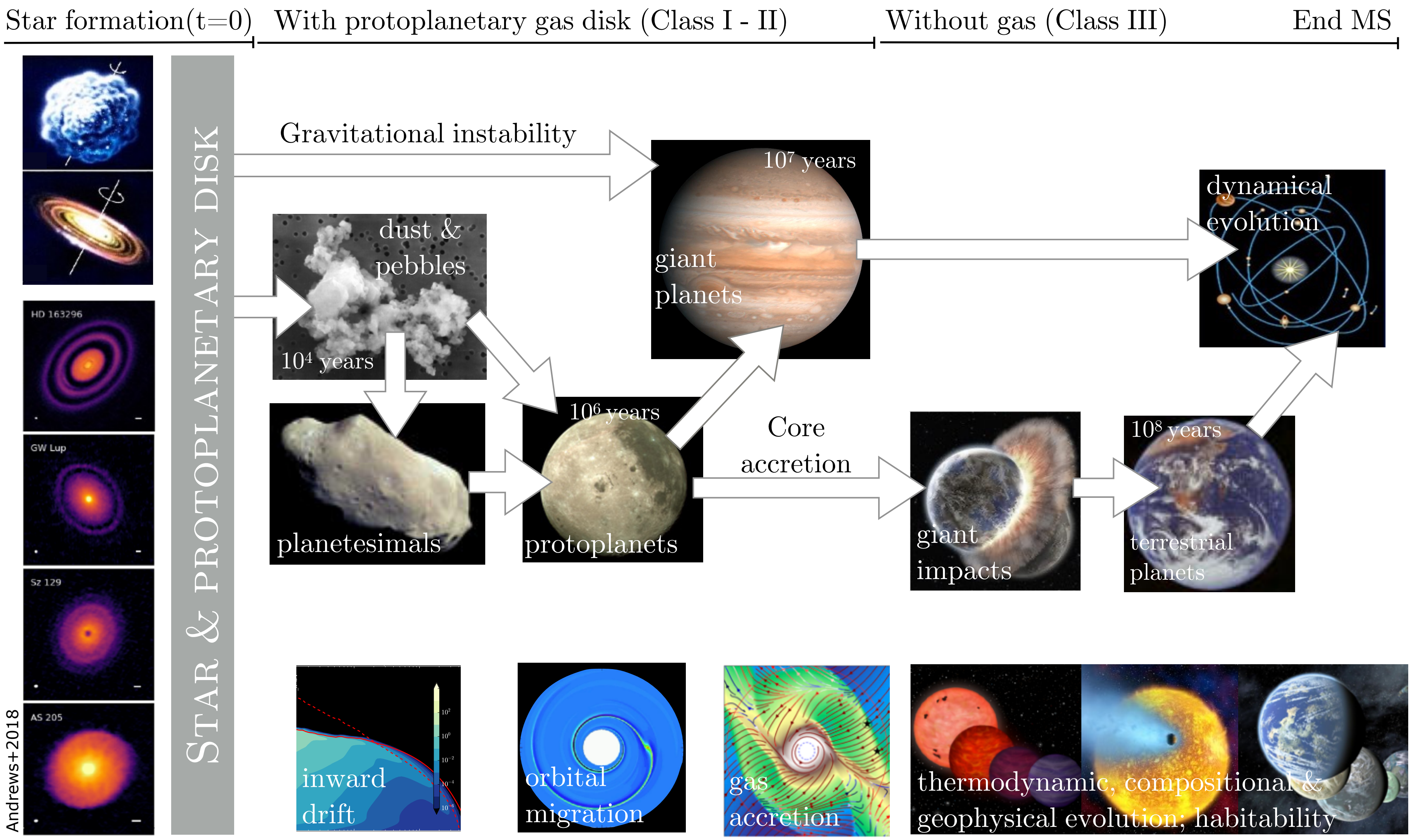
But at least for some, it can give us clues about possible mechanisms responsible for them.

Combine constraints from all major exoplanet observation methods plus Solar System and protoplantary discs



2.
Population synthesis method

Planet formation: stages and physical processes

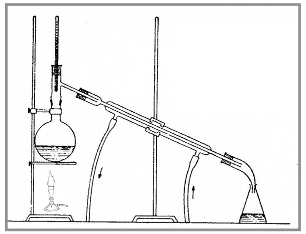


Shifts of paradigms

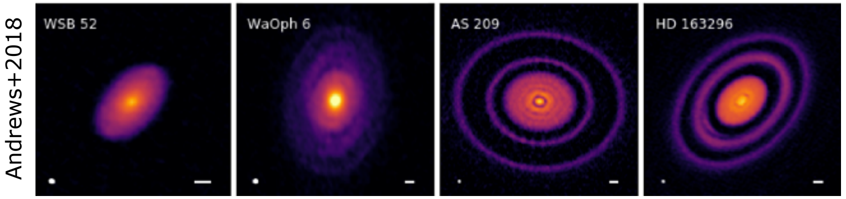
(relative to original Solar System formation theory)

- Mobility (both at pebble and protoplanet level)
- Solid accretion (pebbles, planetesimals, giant impacts)
- Disks: structures, MHD-winds, open system (link to star formation)

How can we test models for all these processes observationally?



The principle

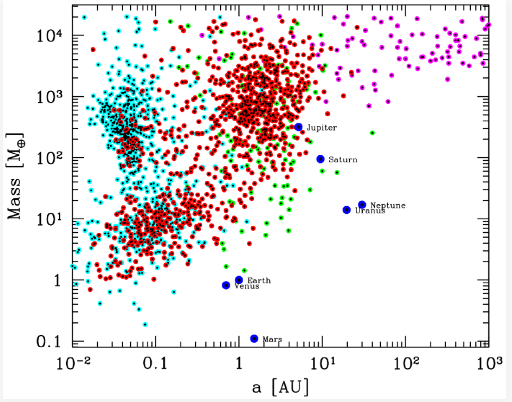


Models of individual processes
Accretion, migration, interiors, …

→ **Global end-to-end formation & evolution model**
Link disk properties ⇒ planet properties

Initial Conditions: Probability distributions of disk properties
Disk gas mass
Disk dust mass
Disk lifetime
From observations

Draw and compute synthetic planet population



Observed population

Building Instrumentation
most compelling future observations

Apply observational detection bias

Predictions
(going back to the full synthetic population)

Stat. Comparison:
Observable sub-population
- Frequencies
- Orbits, masses, radii, luminosities
- Architecture, multiplicity
- Correlations

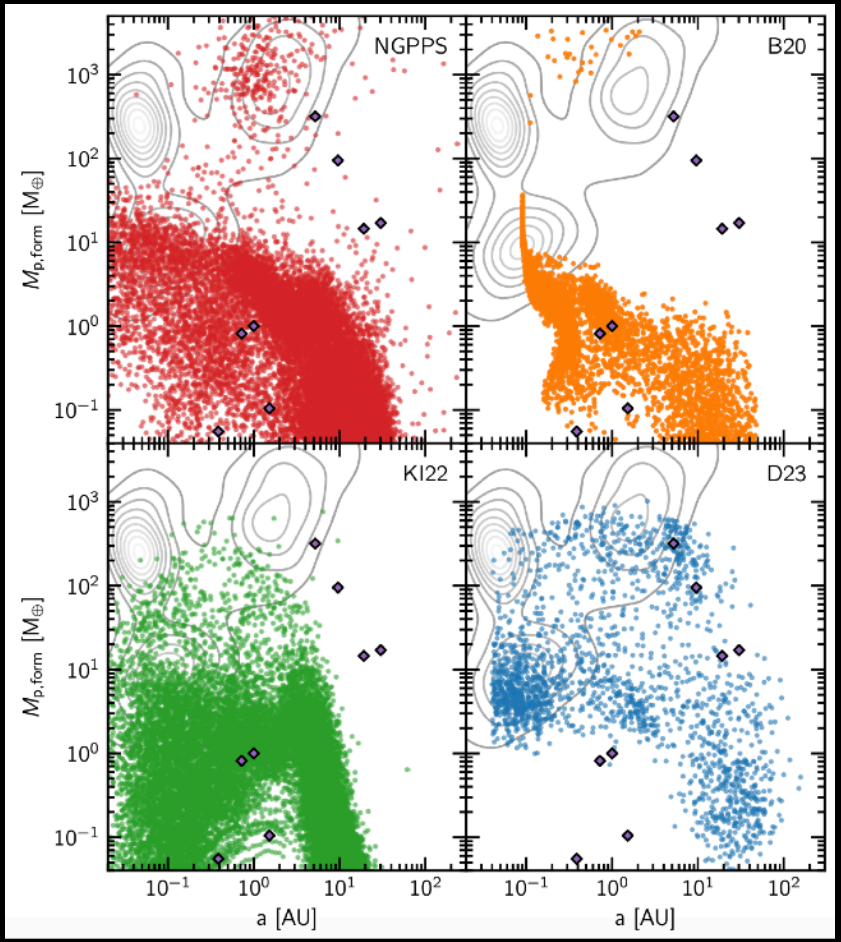
…..

No match: improve, change parameters, new approaches

Match

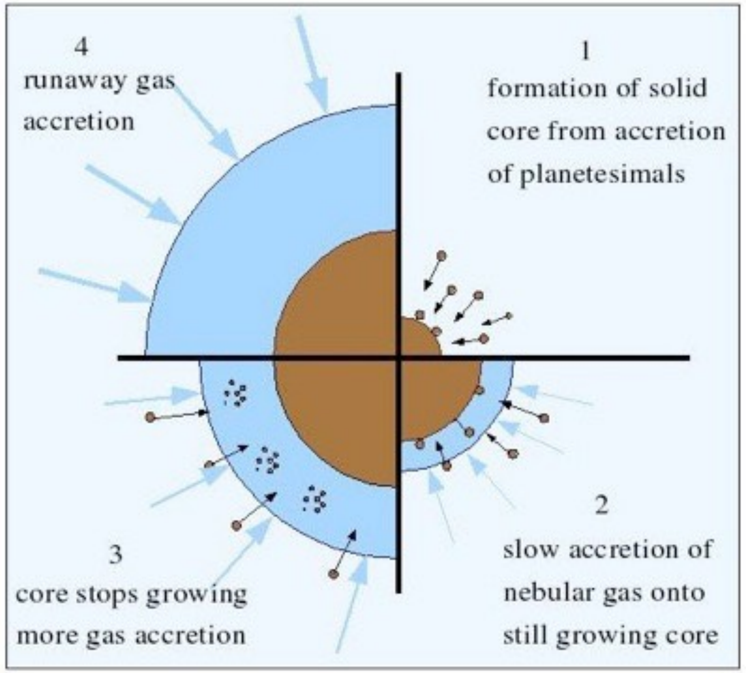
Model solution found

Ida & Lin 2004, Benz+2014, Burn & Mordasini 2025



3.
Global models

Global models



Core accretion paradigm

Perri & Cameron 1974, Mizuno et al. 1978, Mizuno 1980, Bodenheimer & Pollack 1986, Pollack et al. 1996,…

1)Build up critical core
2)Accrete gas

Minimal set of physical processes to be considered

1. Evolution of the protoplanetary disk
2. Accretion of solids
3. Accretion of gaseous envelope (H/He)
4. Orbital migration
5. N-body interaction among (proto)planets

Similar timescales

Feedbacks

An early example

Computer Simulation of the Formation of Planetary Systems

STEPHEN H. DOLE

The Rand Corporation
Santa Monica, California 90406

Received January 25, 1970; revised July 30, 1970

One of the many hypotheses about the formation of the solar system postulates that the planets were formed by the aggregation of particulate matter within a cloud of dust and gas surrounding the newly-formed sun. A test of the validity of one version of this hypothesis was obtained in a computerized Monte Carlo simulation of the process.

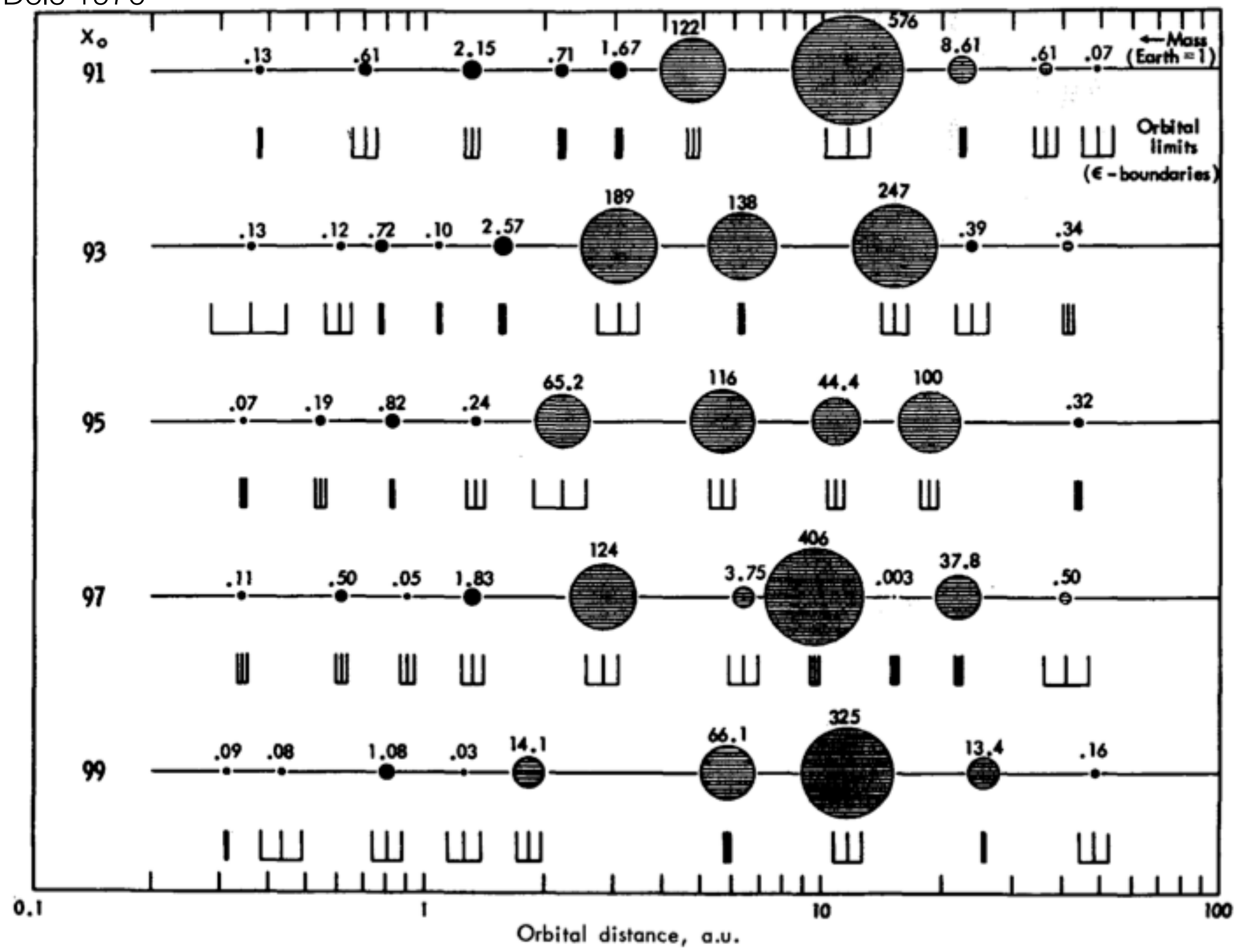
In the model used, nuclei are "injected" into the cloud one at a time, on elliptical orbits. The dimensions of the semimajor axis and the eccentricity of the orbit of each nucleus are determined by using random numbers. As the nuclei orbit within the cloud they grow by aggregation and gradually sweep out dust-free annular lanes. If they grow larger than a specified critical mass they can begin to accumulate gas from the cloud as well. If the orbit of a planet comes inside a certain interaction distance from a planet that was formed earlier, or if the orbits cross one another, the two bodies coalesce to form a single, more massive planet which may then continue to grow by aggregation. The process of injecting nuclei is continued until all the dust has been swept from the system. At this point the run is terminated and the machine output displays the masses and orbital parameters of the planets remaining in the final configuration.

Each planetary system produced by using a different random number sequence is unique. However, all the systems so produced share the major regular features of our solar system. The orbital spacings have patterns of regularity suggestive of "Bode's law." The innermost planets are small rocky bodies; the midrange planets are large gaseous bodies; the outermost planets are generally small. The general pattern of planetary mass distribution is similar to that in our solar system with masses ranging from less than that of Mercury to greater than Jupiter's.

Based on nebular hypothesis and core accretion paradigm: first accretion of solid cores, then accretion of gas if sufficiently massive. In situ.

“Monte carlo computer synthesis”

Dole 1970



-Solar System-like architectures with ~uniform spacing in log

-no close-in planets a<0.1 AU, no distant giant planets

-could clearly not reproduce the exoplanet demographics

~isolation mass for solid accretion

~critical mass for gas accretion

Pre-viscous-accretion disk theory (Lynden-Bell & Pringle 1974)

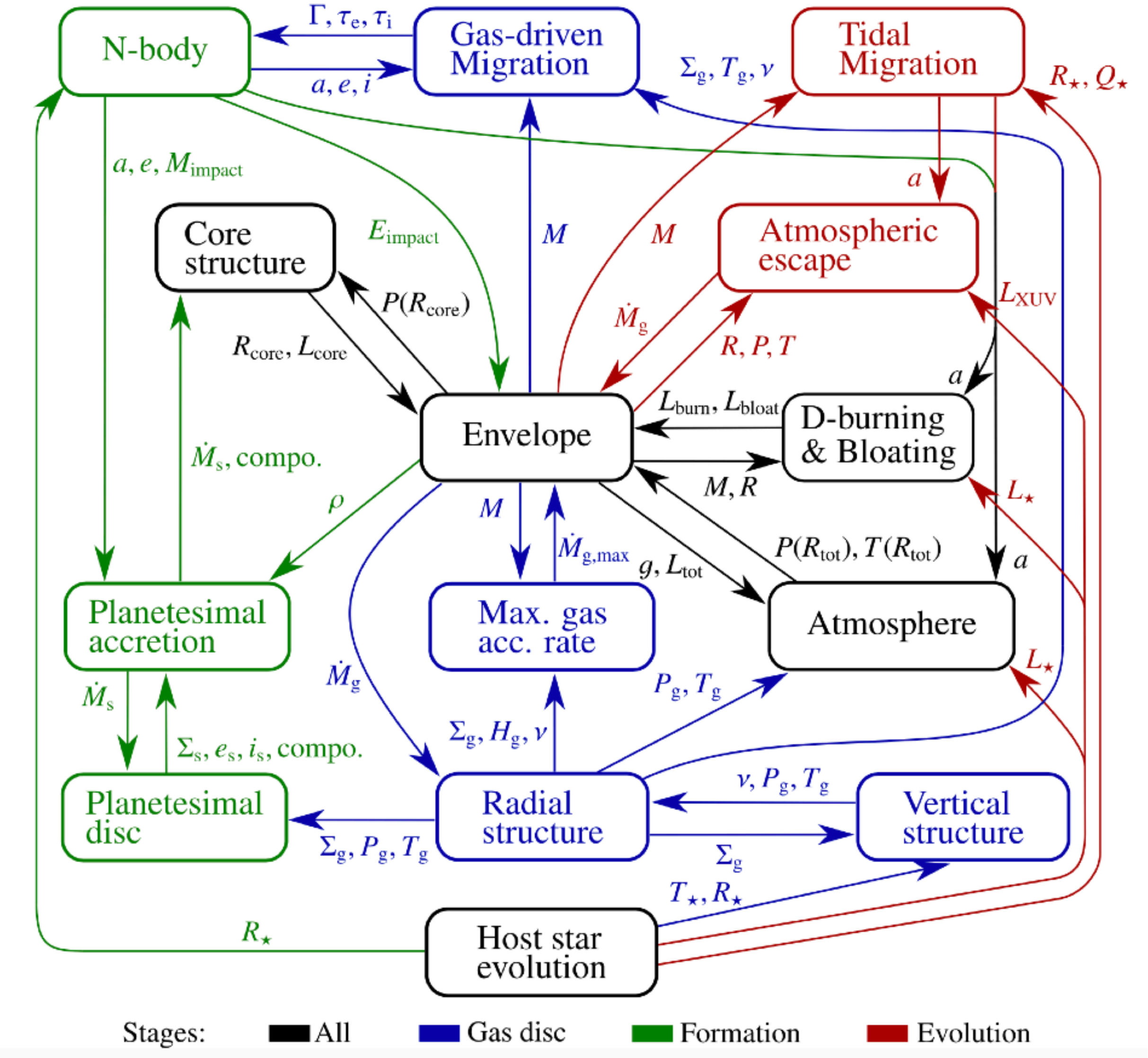
Pre-planetesimal accretion theory (Safronov 1972)

Pre-1D planetary structure theory (Mizuno 1978)

Pre-orbital migration theory (Goldreich & Tremaine 1979)

Reliance of global models on models for specific processes … and on observations!

A modern example: the Generation III Bern Model



Main publications: Alibert et al. 2005, 2013; Mordasini et al. 2009, 2012; Benz et al. 2014; Emsenhuber et al. 2021, 2023

Core accretion paradigm

The Generation III Bern model: a comprehensive end-to-end models of planetary system formation and evolution

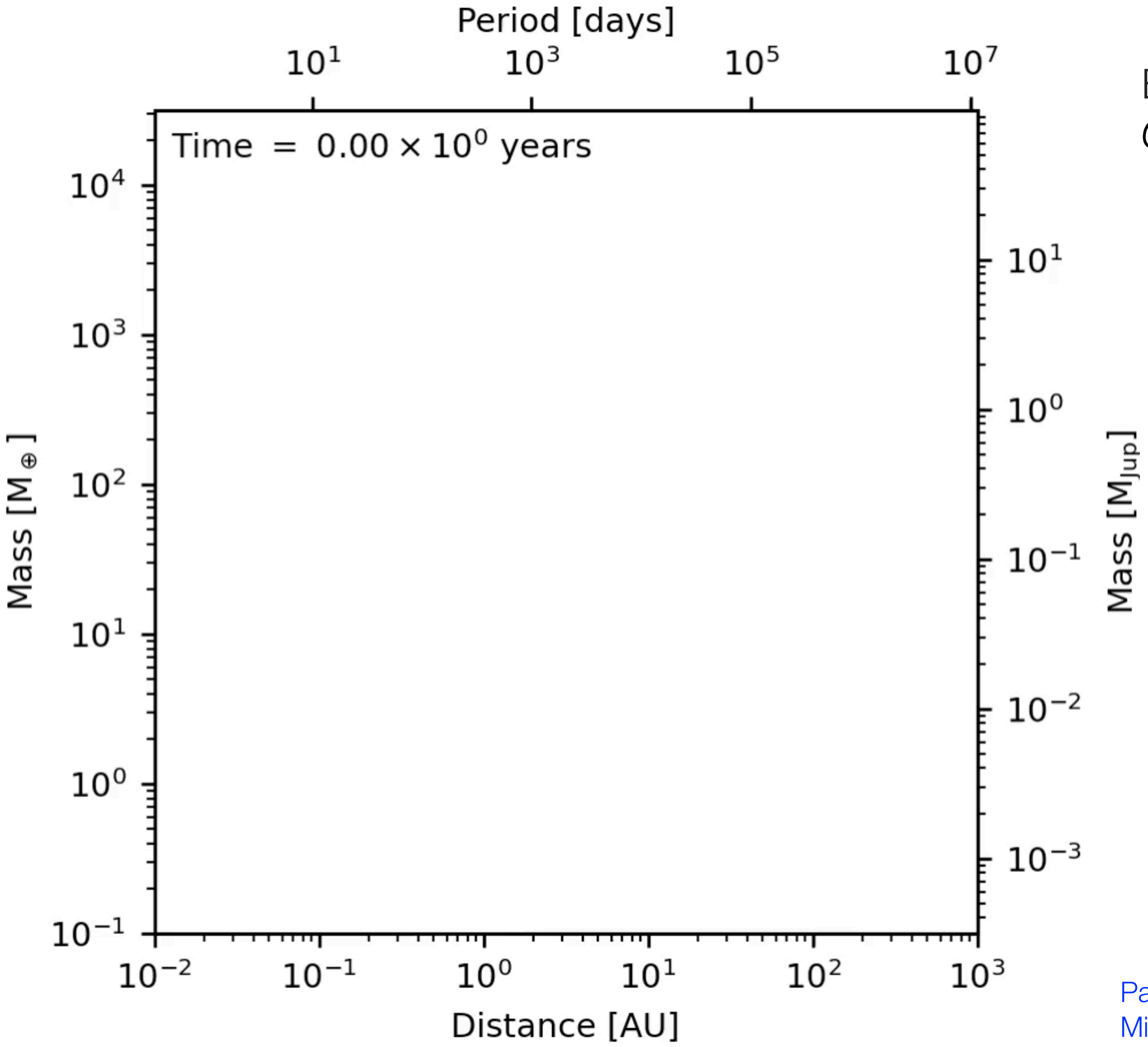
It combines

- *formation* (disk evolution, accretion of gas and solids, orbital migration, N-body interactions, internal structure calculation; first ~100 Myr)
- *long-term evolution* (internal structure, atmospheric escape, tides; following ~5 Gyrs).

Coupled physical processes - modelled by direct solution of governing differential equations, but low-dimensional approximation (axisymmetric disk; spherical planets; only N-body is 3 dimensional).

Direct prediction of all important directly observable quantities for (exo)planets (orb. elements; mass, radius, magnitudes).

NGPPS population synthesis 1 $M_\odot$



Bern Gen III model used to generate the *New Generation Planetary Population Synthesis, NGPPS*

- 1000 systems (stars); solar-like host stars
- 100 initial embryos per system of 1 M_{luna}
- 4 Monte Carlo variables: [M/H], initial disk gas mass, external photoevaporation rate, disk inner edge

Gas-dominated

silicate/iron

with volatiles

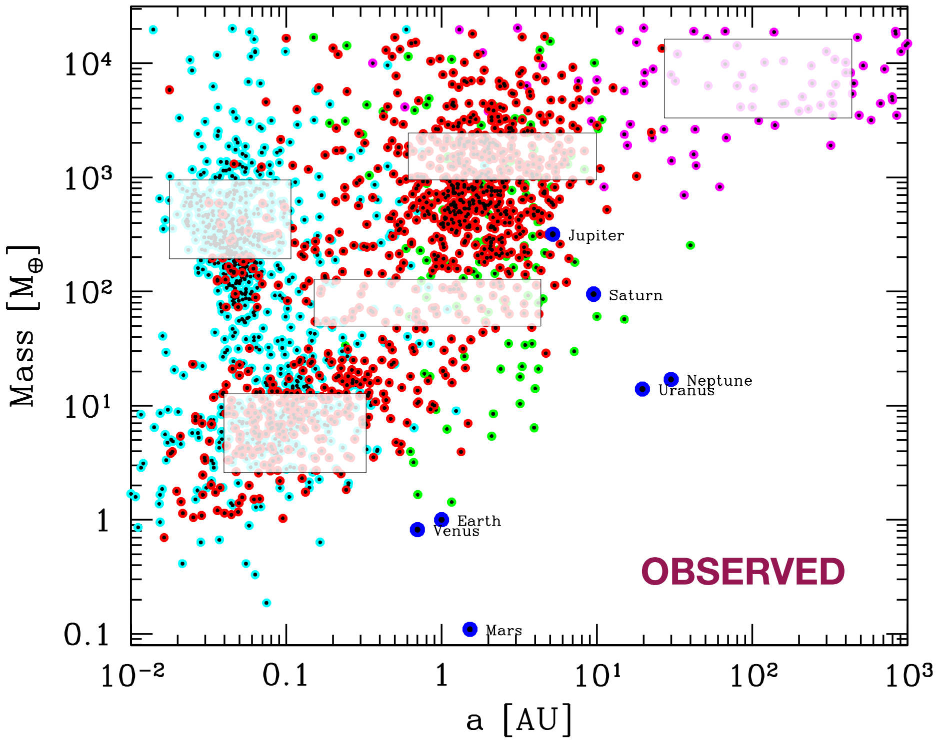
Lost

Eccentricity

Solar system

Papers: NGPPS I-VIII: Emsenhuber+2021a,b; Schlecker+2021a,b, Burn+2021, Mishra+2021, Di-Chang+2025 plus Emsenhuber+2023, Burn+2024, …

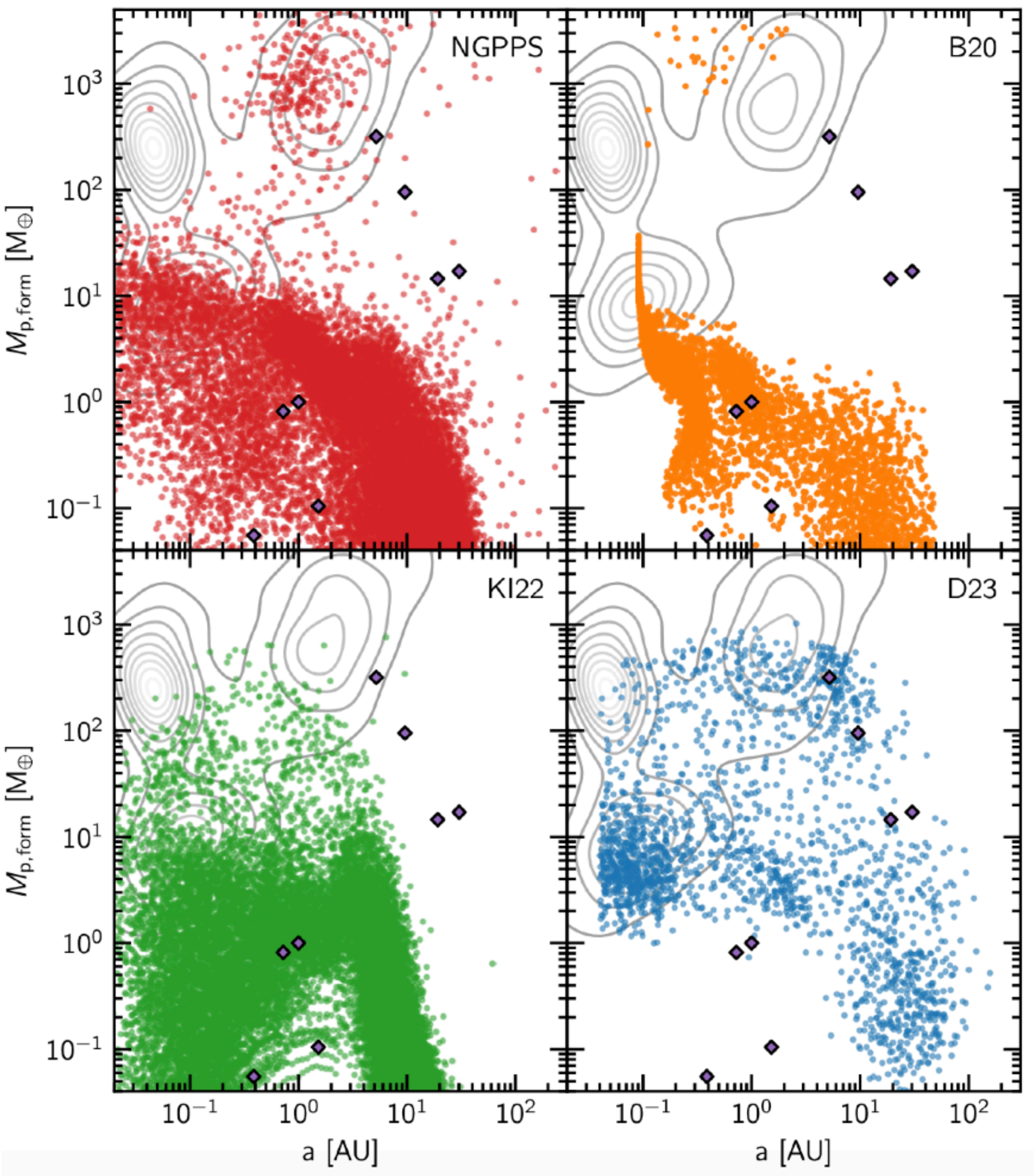
Overall result



Diversity of initial conditions (disk properties) leads to diversity of planetary systems similar as observed.

But how well does it compare quantitatively to observations?

- radial velocities (Emsenhuber+acc., NGPPS VII)
- Kepler (Mulders+2019, Burn+2024, Dichang+2024)
- SPHERE SHINE (Vigan+2020)
- Microlensing (Suzuki+2018, Zang+2025)

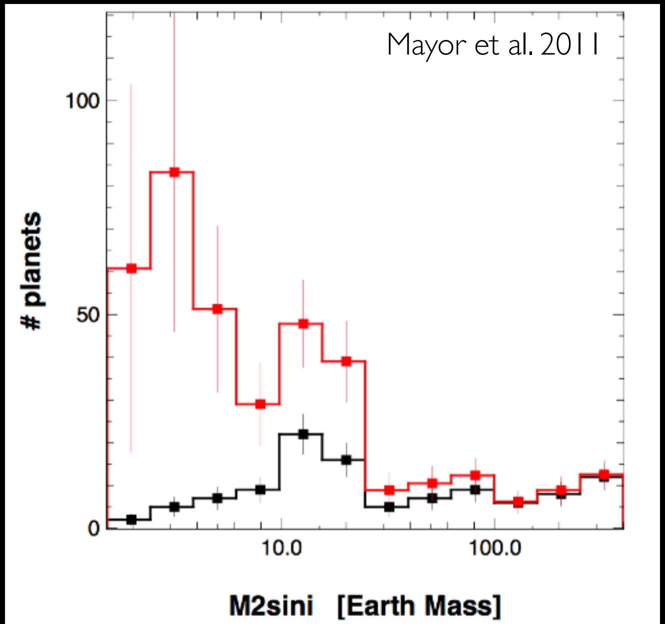


Comparison of pop. synth models in Burn & Mordasini (2024)

- NGPPS (Emsenhuber+2021)
- Bruegger+2020
- Kimura & Ikoma 2022
- Drazkowska+2023

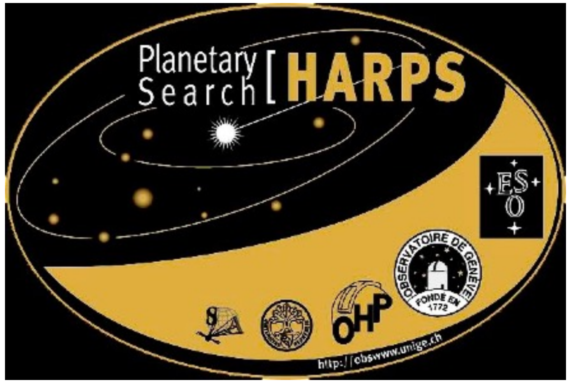
See also models of

- Bitsch+2015, 2022
- Chambers 2018
- Alessi & Pudritz 2022
- Pan, Liu+2024



4.
Comparisons with observations

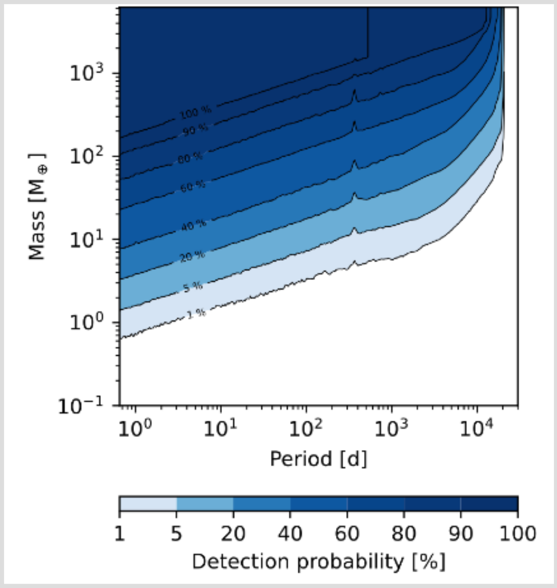
Radial velocities I: The HARPS GTO survey



See also results of California Legacy Survey (Rosenthal+2021, Fulton+2021, ….)

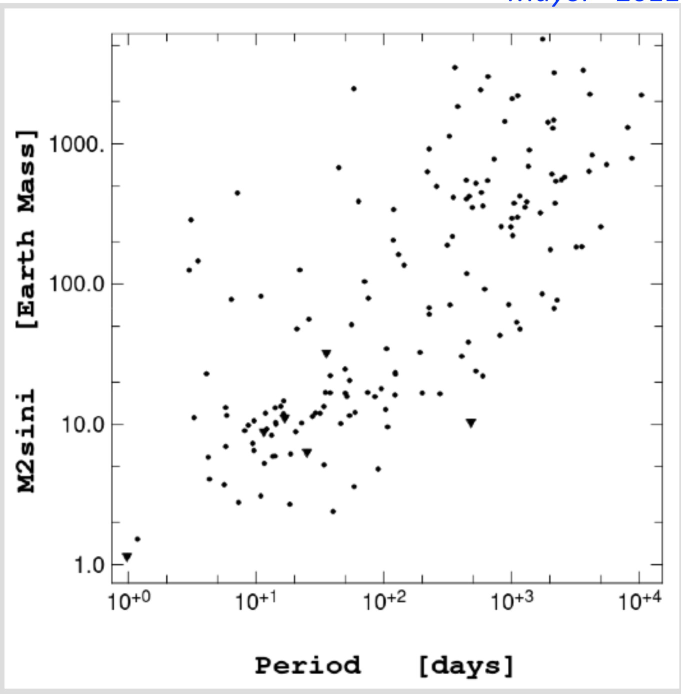
- HARPS: High Accuracy Radial velocity planet searcher (spectrograph for radial velocity measurements) at ESO 3p6 (Mayor et al. 2003)
- RV accuracy of ~1 m/s: first detection of low-mass exoplanets
- Volume-limited sample within 50 pc with 822 low-activity solar-like stars
- Combined w. Coralie survey for longer baseline (Udry et al. 2000)
- Known mean detection bias from signal injection and recovery tests
- Statistical analysis in Mayor+2011

Mayor+ 2011



Importance of large surveys with well-defined bias (like Kepler or direct imaging surveys; in future PLATO, NIRPS GTO, GAIA, Roman Space Telescope, …)

Mayor+ 2011



Nb of discovered exoplanets: 161

Radial velocities III: comparison mass-distance diagram

- Same approach as Mordasini+2009: "observe" NGPPS with HARPS GTO
- Draw randomly 882 synthetic systems out of 1000 NGPPS systems
- Draw inclination assuming random orientations of systems to get system sin(i)
- Include effects of inter-planet inclinations (model output)
- Apply mean detection bias from Mayor+2011

Observed (Mayor+2011)

Nb of planets: 161

Nb of stars w. planets: 102

Mean obs. multiplicity: 1.58

Synthetic biased

Nb of planets: 294^{+30}_{-27}

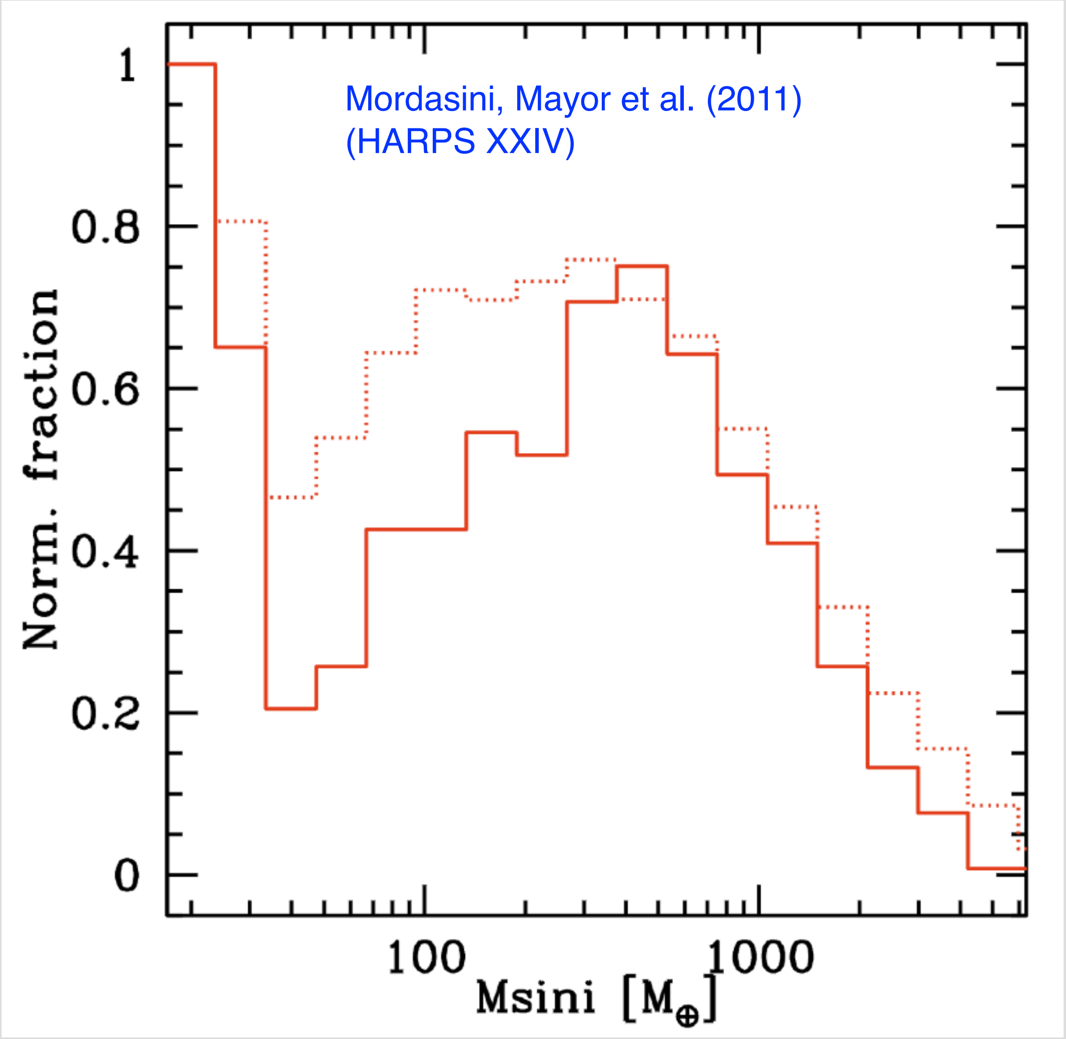
Nb of stars w. planets: 200^{+18}_{-17}

Mean obs. multiplicity: $1.47^{+0.3}_{-0.25}$

- Agreement: similar global structure: relative distribution (concentrations, voids)
- Agreement: Mean multiplicity ⇒ system architect.
- Disagreement: Factor ~2 in absolute number. Poss. explanations: Initial conditions? Cluster environment (cf. Winter+2020)? Not optimised…
- Disagreement: Hot Jupiters. Kozai plus tidal circularisation missing in model. There are eccentric proto-Jupiters. Disk-dominated Type 2 migration rates?

Adapted from Emsenhuber, Mordasini, Mayor, Marmier, Udry et al. in press (NGPPS VII)

Radial velocities IV: Planetary mass function (distribution of Msini)



Mass distribution actual and synthetic detected planets

- Agreement: Fundamental bimodal structure
- Agreement: Change in regime at ~30 $M_\oplus$: smoking gun of core accretion: runaway gas accretion M_{core}~M_{enve}~15 $M_\oplus$ (see also Bennet+2021, but also Bertaux & Ivanova 2022, Zang et al. 2025).
- Disagreement: Giant planets too massive (~400 vs ~700 $M_\oplus$) and too numerous by about 45%.
- Disagreement: Too few intermediate mass planets by factor of ~60% (planetary desert, Ida & Lin 2004, Mayor & Udry 2008, Bouchy+2009). Can observationally constrain gas accretion rate in runaway and disk-limited phase.

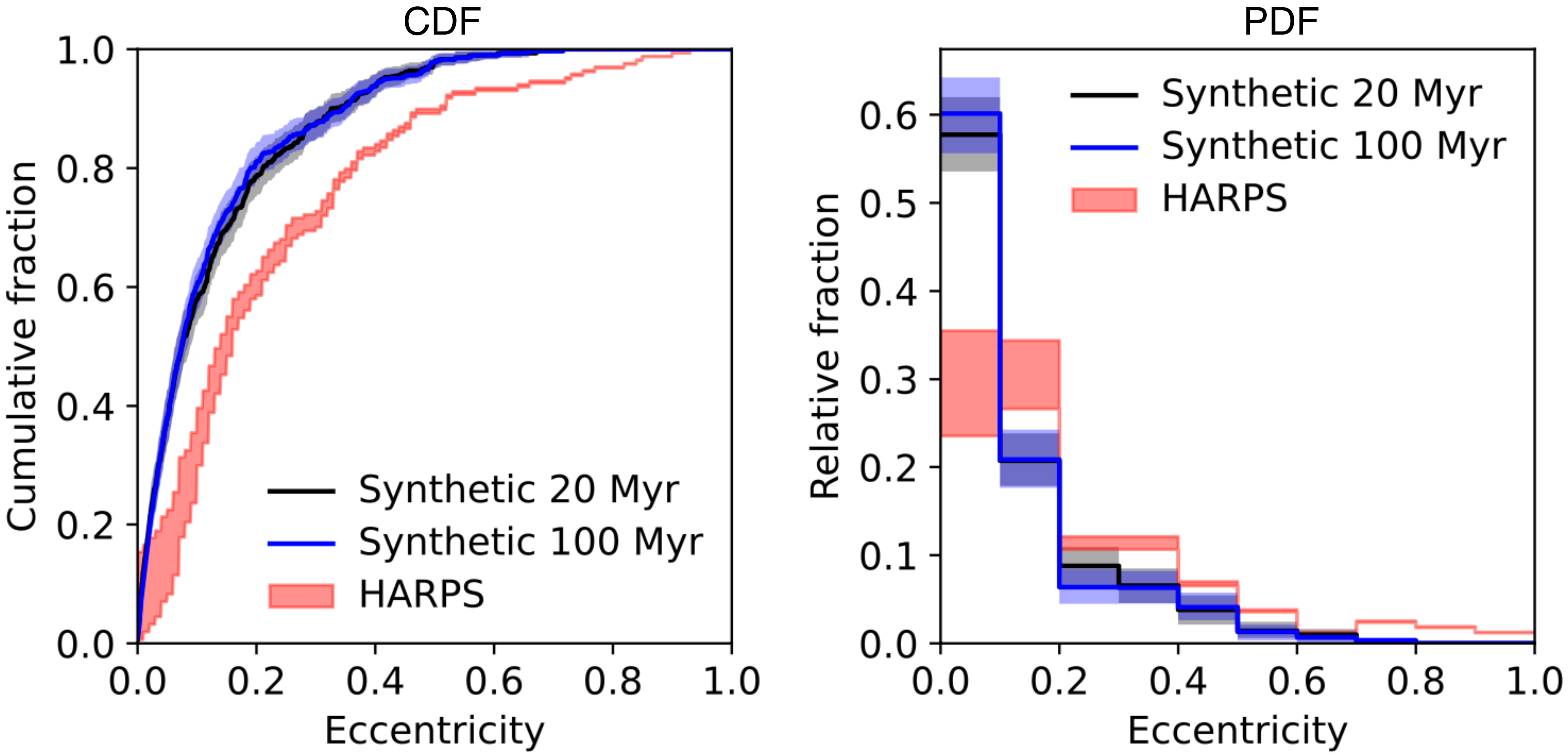
⇒ too fast and too long gas accretion (cf. Nayakshin+2019).

Similar for gas accretion rate derived from several 3D hydrodynamic models (Machida+2010, Bodenheimer+2013, Choksi+2023…)

Possible explanations: low viscosity disks (Ginzburg & Chiang 2019) with efficient gap formation (Aoyama & Bai 2023), magnetic regulation (Batygin 2018, Cridland 2018), angular momentum barrier (Takata & Stevenson 1996), 3D circulation (Szulagyi et al. 2014), …

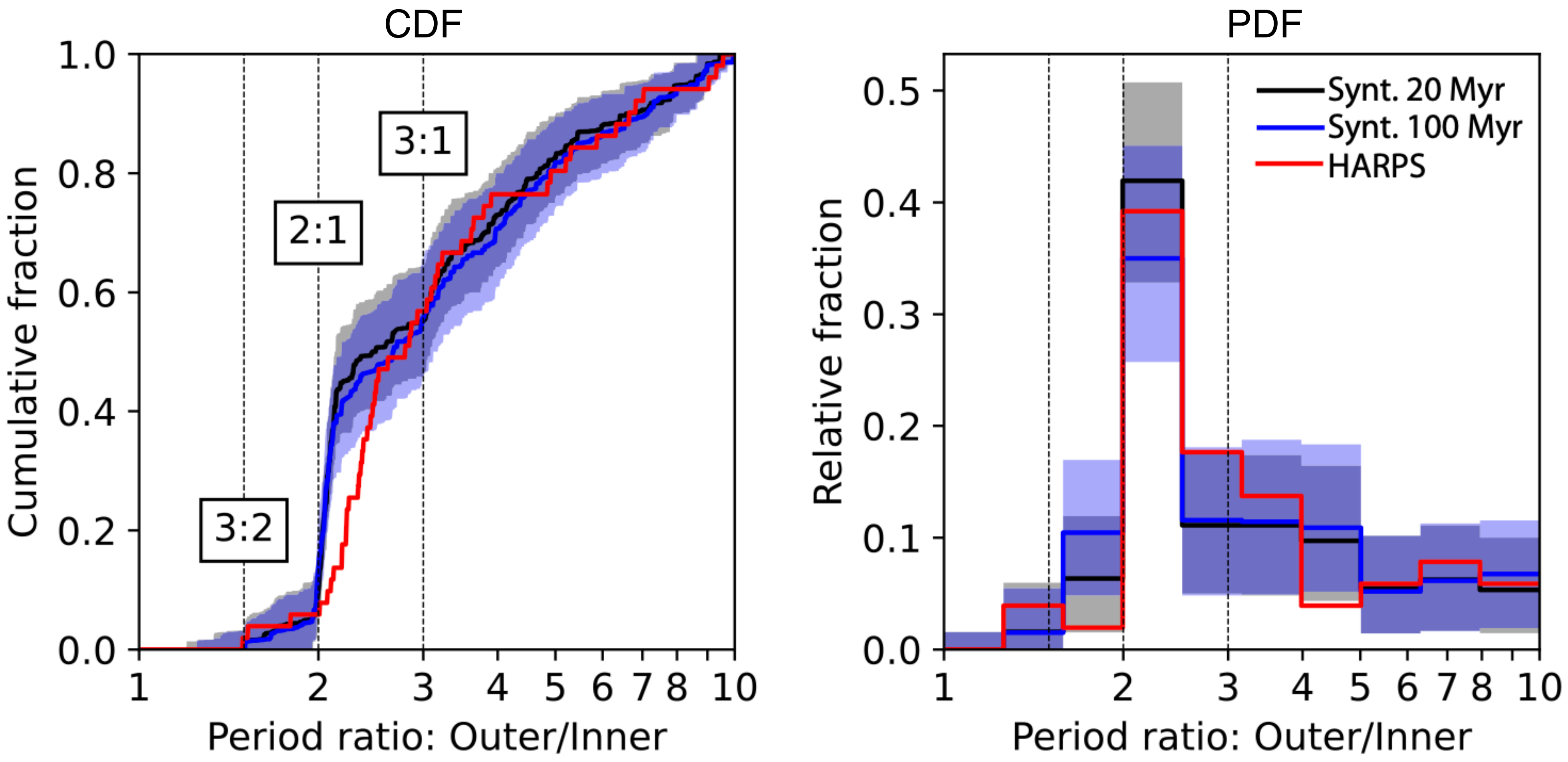
Population synthesis makes it possible to *quantify* discrepancies between theory and observations.

Radial velocities V: eccentricities



- Agreements: good fit at intermediate values (planet-planet scattering, cf. Juric & Tremaine 2008; Ford & Rasio 2008)
- Disagreement: offset by about 0.07 towards higher eccentricities in HARPS relative to synthetic population
- Disagreement: no very high e > 0.7 planets in synthesis
  - Model: Too strong damping? Other mechanisms increasing e than planet-planet scattering? External perturbers / binaries ? Excitation by gas disk for massive planets (Kley & Dirksen 2006)
  - Observations: overestimation of e (known bias of RV method, Lucy & Sweeney 1971, Hara+2019)

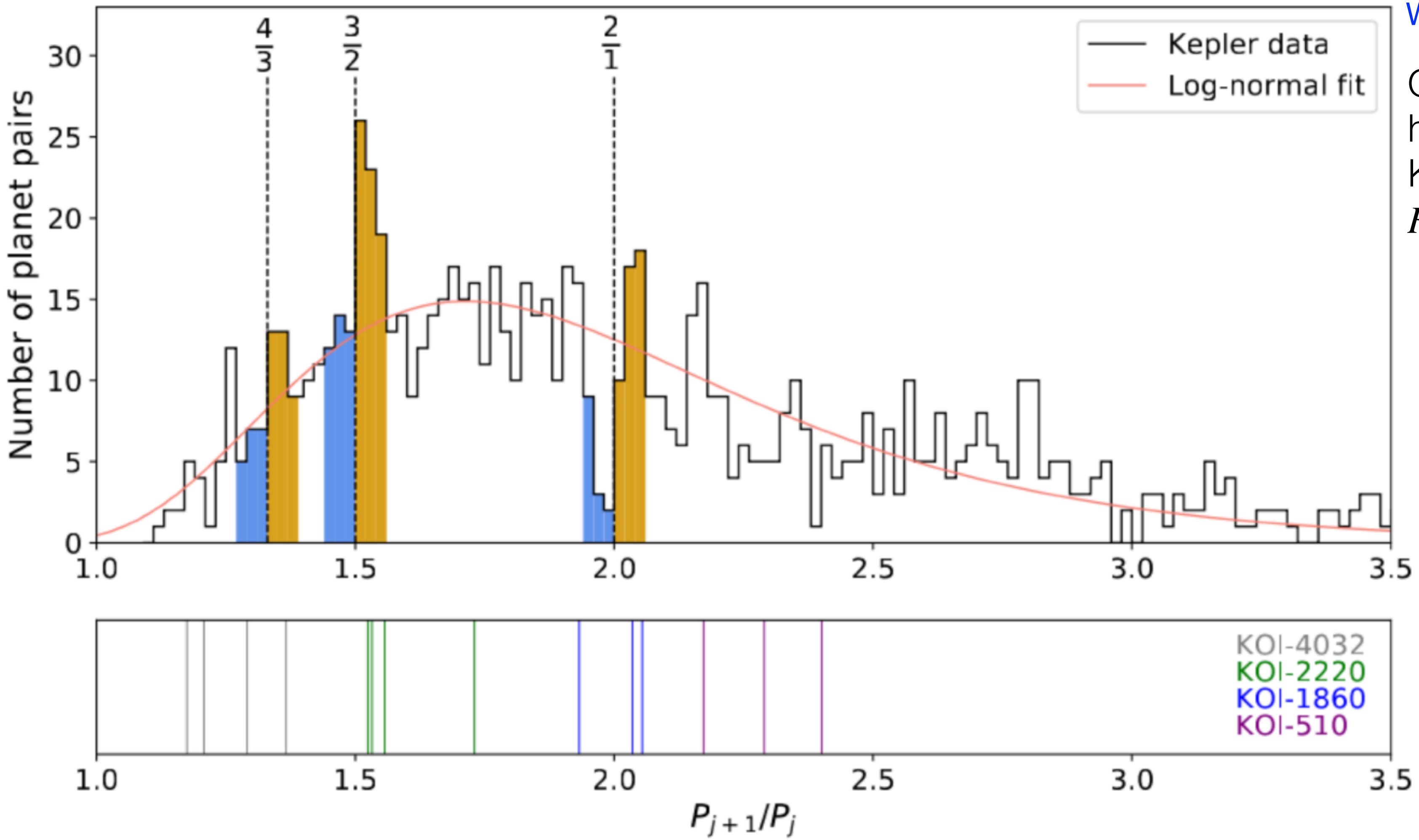
Radial velocities VI: period ratio of adjacent planet pairs



- Agreement: good agreement of shape in most parts, range, both high and low ratios
- Disagreement: too many very close to / in 2:1 MMR in synth. pop. compared to observations
  - Model: Stochastic migration? Too much orbital migration overall? Effect of tides (missing in model)?
  - Observations: two planets in 2:1 interpreted as single eccentric planet?

Comparison with Kepler survey I: period ratios of adjacent planet pairs

Take advantage that model also predicts radii. Cross-compare with Kepler observations.



Weiss+2023 (PPVII)

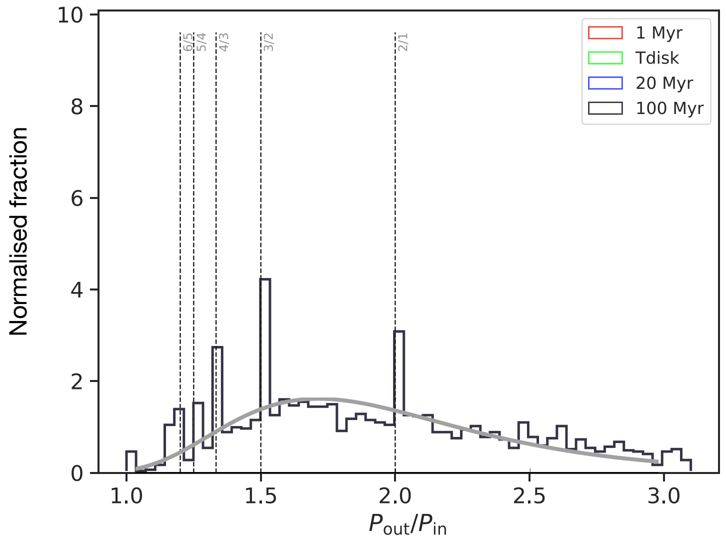
Observed period ratio histogram for small Kepler planets (R<4 $R_\oplus$, short periods)

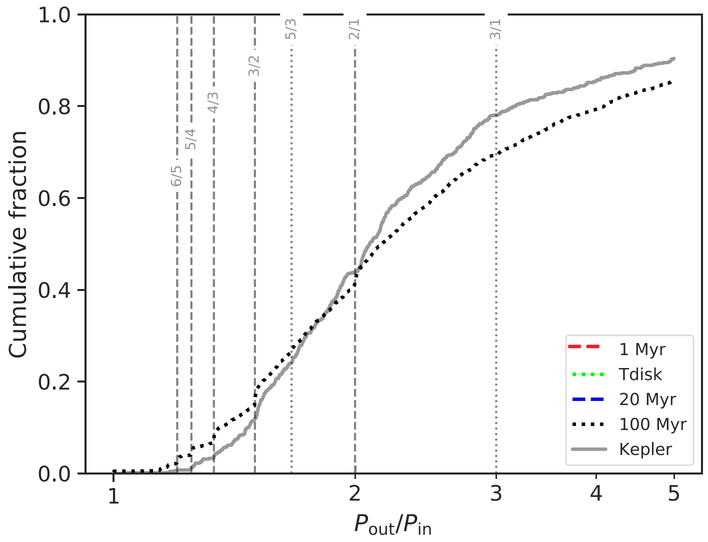
Same data as for comparison with HARPS

See also Mulders+2019

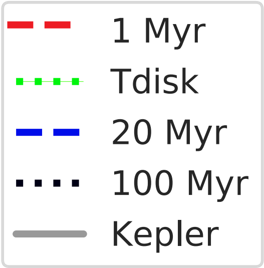
How does this look like in the synthetic population?

Comparison with Kepler survey II

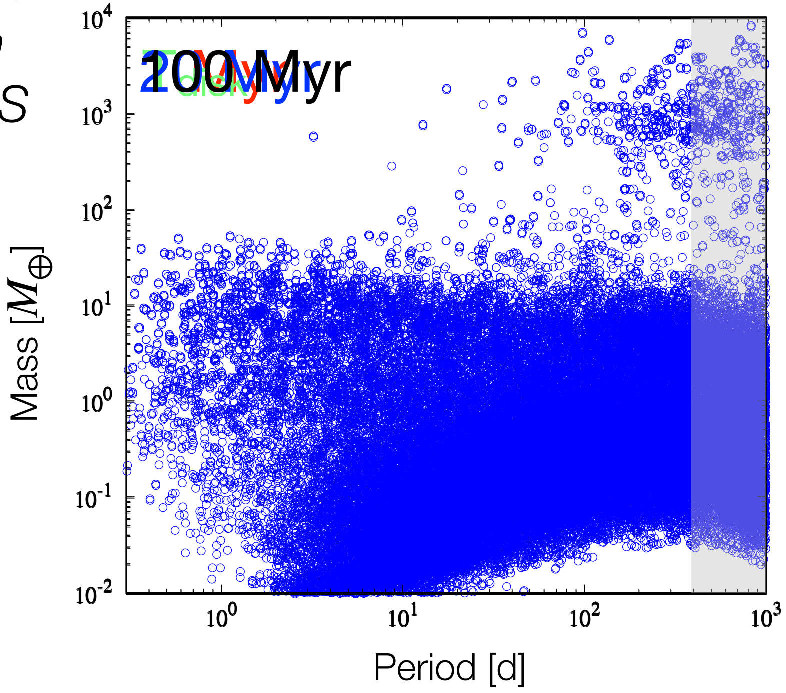




P<400 d



Temporal evolution of period ratio in synthetic NGPPS population



1 Myr (in gas disk): majority $P_{out}/P_{in} < 1.2$. Gas damps eccentricities, stabilises orbits (Kley & Nelson 2012). Tight packing from oligarchic planetesimal growth phase (relative spacing ~10 mutual Hill radii, Kokubo & Ida 1998) and convergent migration.

T_{disk} (~3 Myr): Damping vanishing, the frequency of pairs with $P_{out}/P_{in} < 1.2$ decreases. MMRs like 6/5, 5/4, 4/3 strongly populated. Pairs with $P_{out}/P_{in} > 2$ still rare.

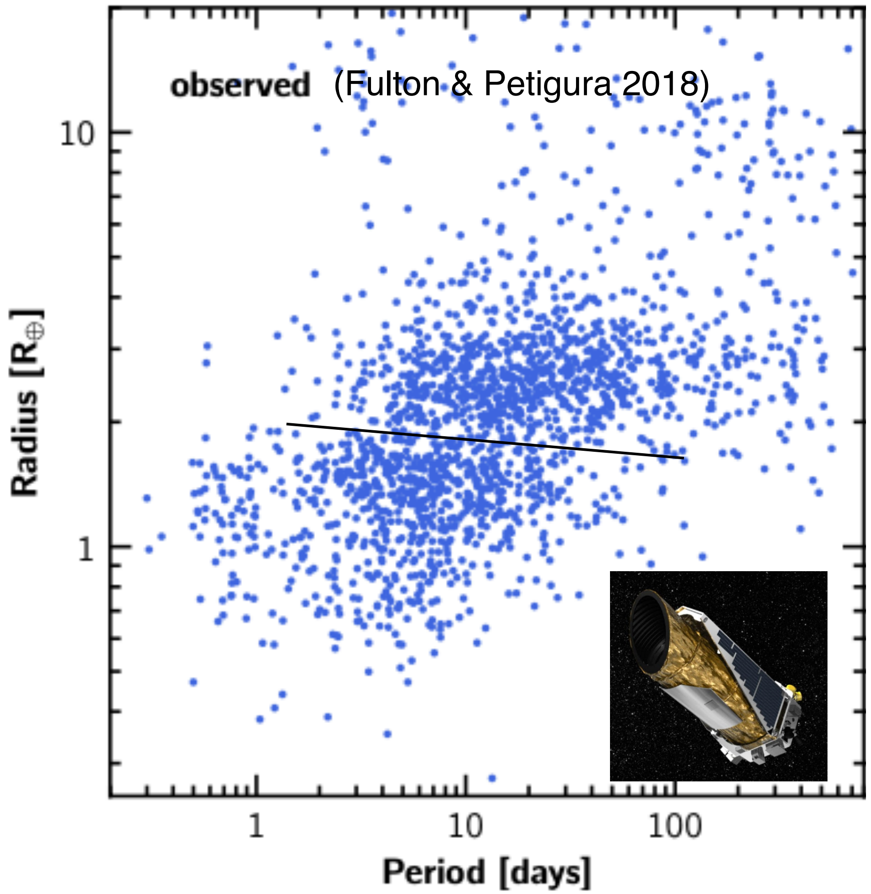
20 Myr: Large change between T_{disk} and 20 Myr: many resonances break, a lot of giant impacts (Ida & Lin 2010; Izidoro et al. 2017)

100 Myr: Between 20 and 100 Myr, fraction of planets in and near the resonances, especially in tighter ones, decreases more. Fraction of resonant systems, however, still larger (by ~60%) than observed in Kepler data (~5 Gyr) (but see Leleu+2024b).

In model: pairs exactly in resonances. No self-consistent tides - N-body coupling included yet in model. See the temporal evolution with PLATO (cf. Dai et al. 2024)?

Comparison with Kepler survey III: the radius valley

Does NGPPS reproduce the radius valley (Fulton gap), one of Kepler's most important results? Not in the original version assuming condensed ice layers in the interior structure model (Owen & Wu 2017, Jin & Mordasini 2018)



Improved NGPPS: AQUA water EOS (Haldemann et al. 2020) with correct phases & mixed H/He + water envelopes

Valley separates larger water-rich migrated sub-Neptunes w. supercritical water/steam envelopes from smaller dry super-Earth (silicate+iron) formed inside of the water iceline

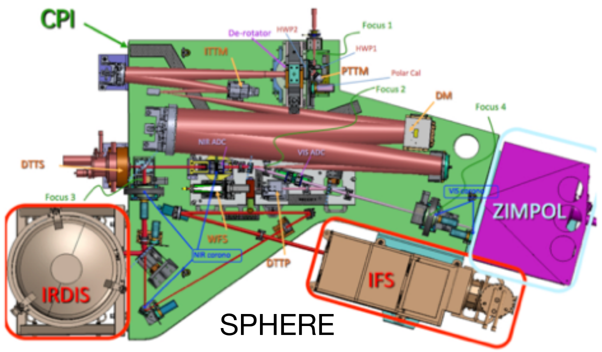
cf. Turbet+2019, Zeng+2019, 2021; Mousis+2020, Venturini+2020a,b, 2024

Comparison with direct imaging

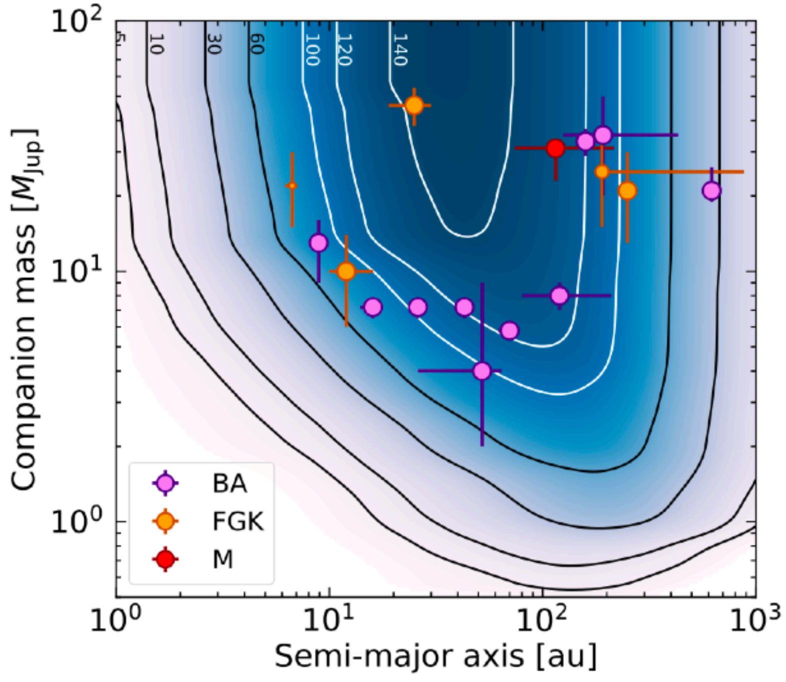
Probes very different kind of planets and a different observable (magnitudes / luminosity)

SPHERE@VLT SHINE GTO survey (Vigan et al. 2020)
150 stars

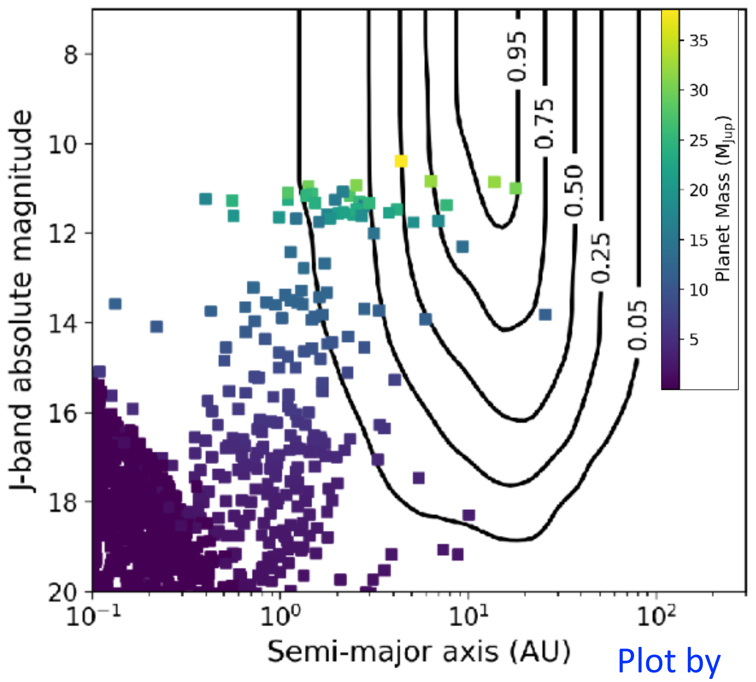
cf. Nielsen et al. 2019 GPIES



Actual detections & sensitivity maps



Synthetic population & sensitivity maps



Plot by Clemence Fontanive

Fraction of FGK stars w. planets ($M=1\text{-}75\ M_J$, $a=5\text{-}300$ AU)

Observed: $5.8^{+4.7}_{-2.8}$ %

Synthetic: $3.4^{+0.5}_{-0.5}$ %

- Agreements: overall frequency, mass-luminosity relation (β Pic b)
- Distant giants in synthesis: Single, massive, eccentric planets from scattering events (see Marleau+2019b), mean eccentricity: 0.39
- Disagreement: No HR 8799-like systems: 4 distant massive giants on ~circular orbits
- Structured disks? Formation by gravitational instability?

Comparison with microlensing: mass distribution

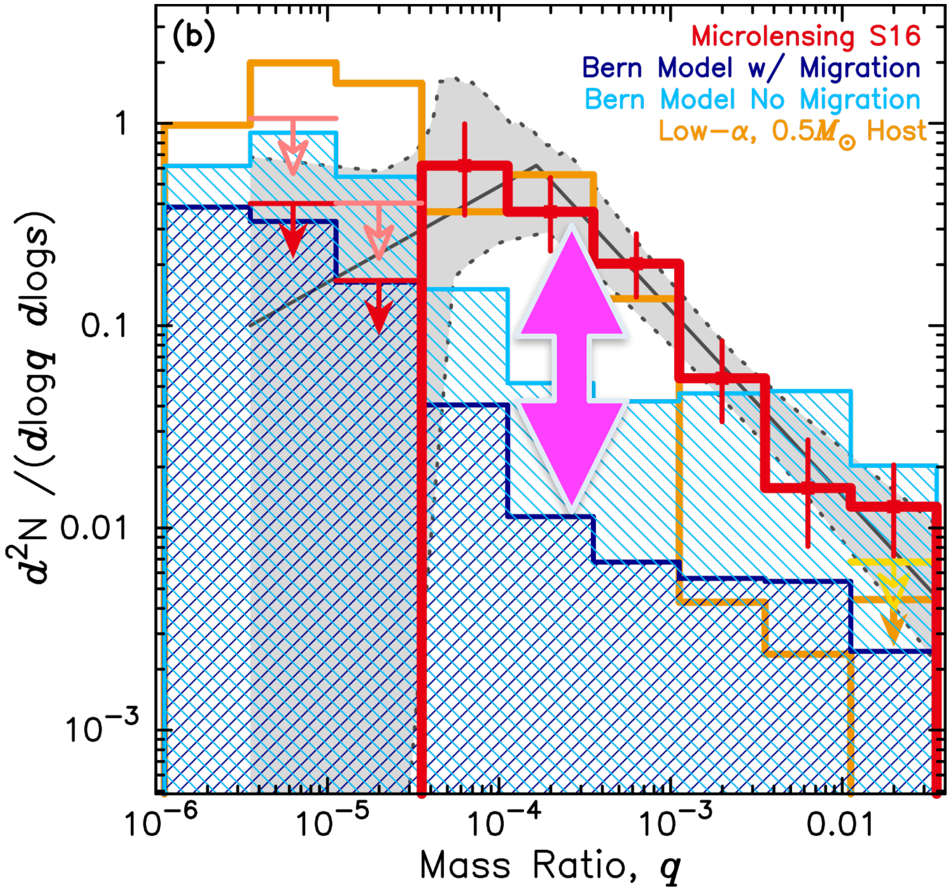
Synthetic mass distribution

Imprints of core accretion are visible in predicted synthetic mass distribution

-the accretion of solids
-the critical core mass
-the accretion of gas
-planetary desert

Observations (MOA)

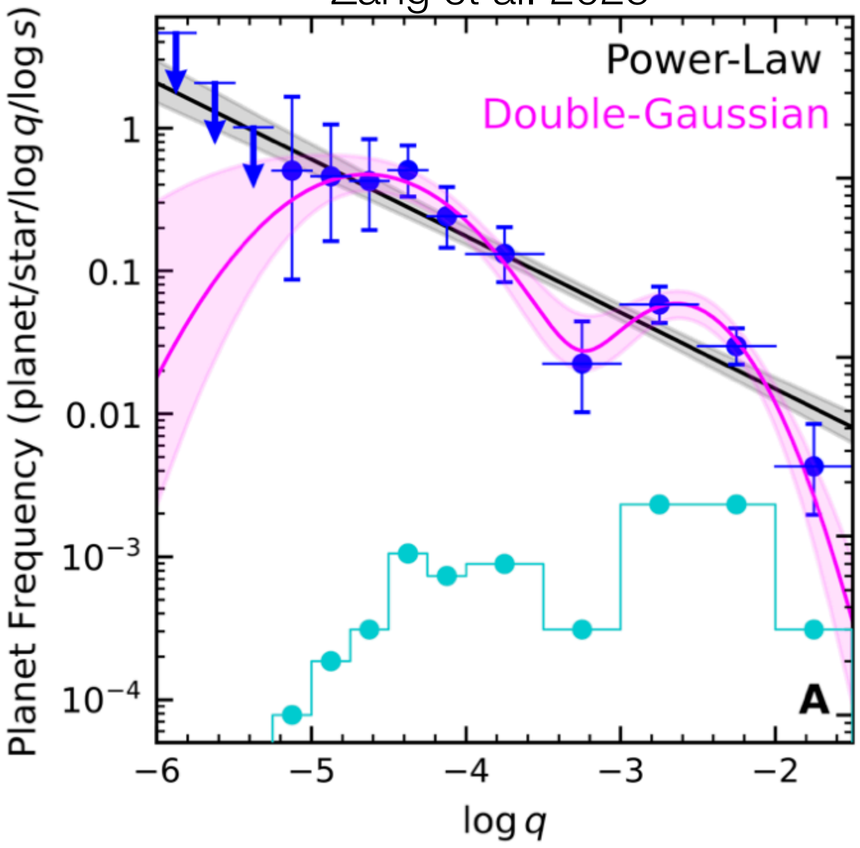
M-dwarfs, beyond iceline, Suzuki et al. 2018



-maximum (?) at ~20-50 M_{Earth}
-much higher frequency than synthetic
-low disk viscosity could help

Observations (KTM Net)

Zang et al. 2025



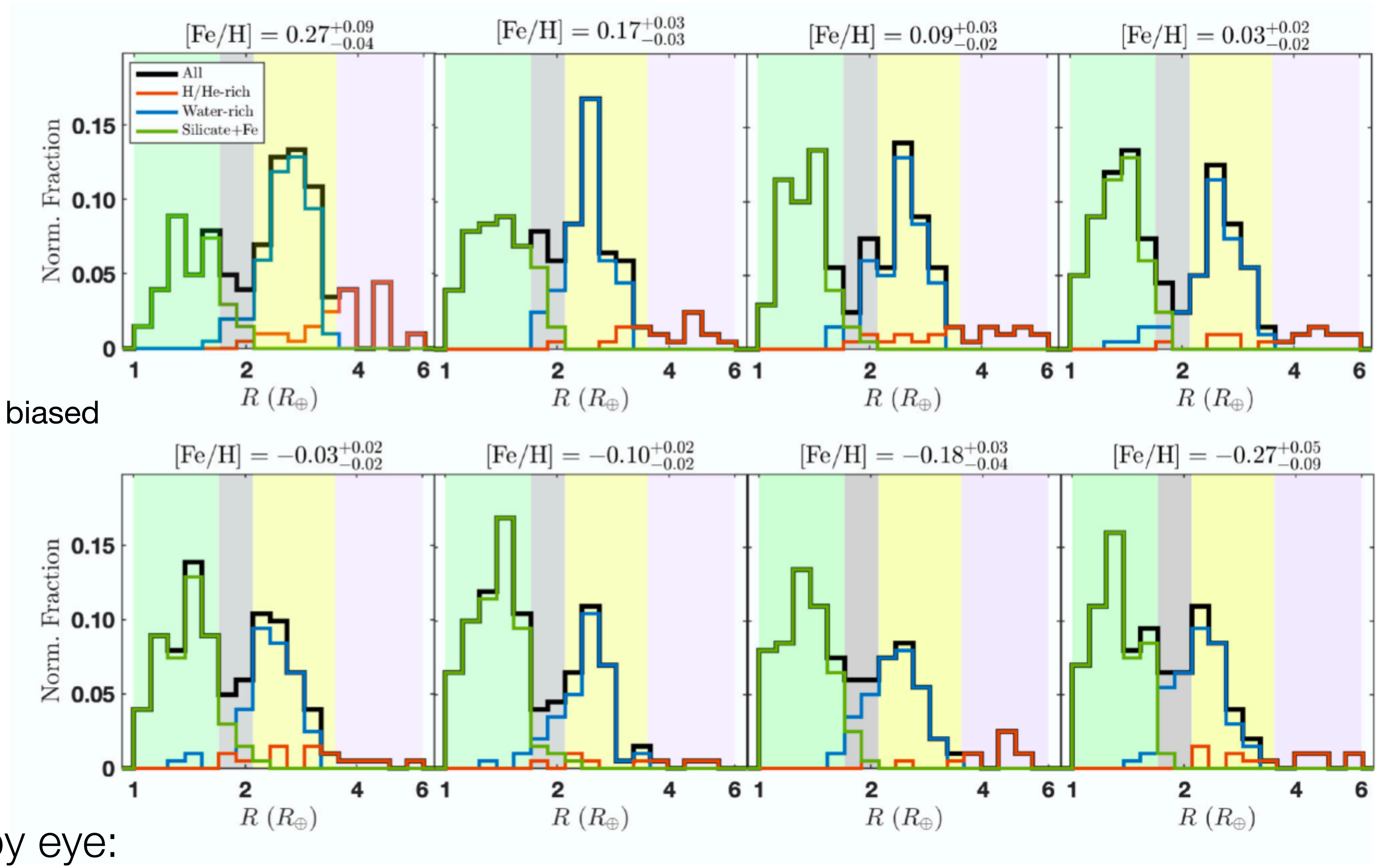
-new data points at a potential minimum
-detailed comparison necessary

Roman ST: a game-changer

Valley morphology: dependency on host star [Fe/H]

Bin synthetic radius distribution according to host star [Fe/H].

synthetic, biased



NGPPS VIII (Chen et al. A&A accepted.)

Visible by eye:

- Ratio super-Earth to sub-Neptunes increases with decreasing [Fe/H]
- Sub-Neptunes become smaller with decreasing [Fe/H]

Quantitative comparison with PAST III (NGPPS VIII)

Comparison with PAST III (Chen et al. 2022) using LAMOST-Gaia-Kepler catalog

Valley morphology quantified by 5 metrics using the number of super Earth SE, sub Neptunes SN, valley planets VP, Neptunian planets NP:

1. The contrast C_{valley} $$C_{\text{valley}} = \frac{N_{\text{SE}} + N_{\text{SN}}}{N_{\text{VP}}}.$$

2. The asymmetry A_{valley} $$A_{\text{valley}} = \log_{10}\left(\frac{N_{\text{SE}}}{N_{\text{SN}}}\right)$$

3. The average (sub)Neptune radius R^{+}_{valley}

$$R^{+}_{\text{valley}} = \frac{1}{N_{\text{SN}} + N_{\text{NP}}} \sum_{i}^{N_{\text{SN}}+N_{\text{NP}}} R_i,$$

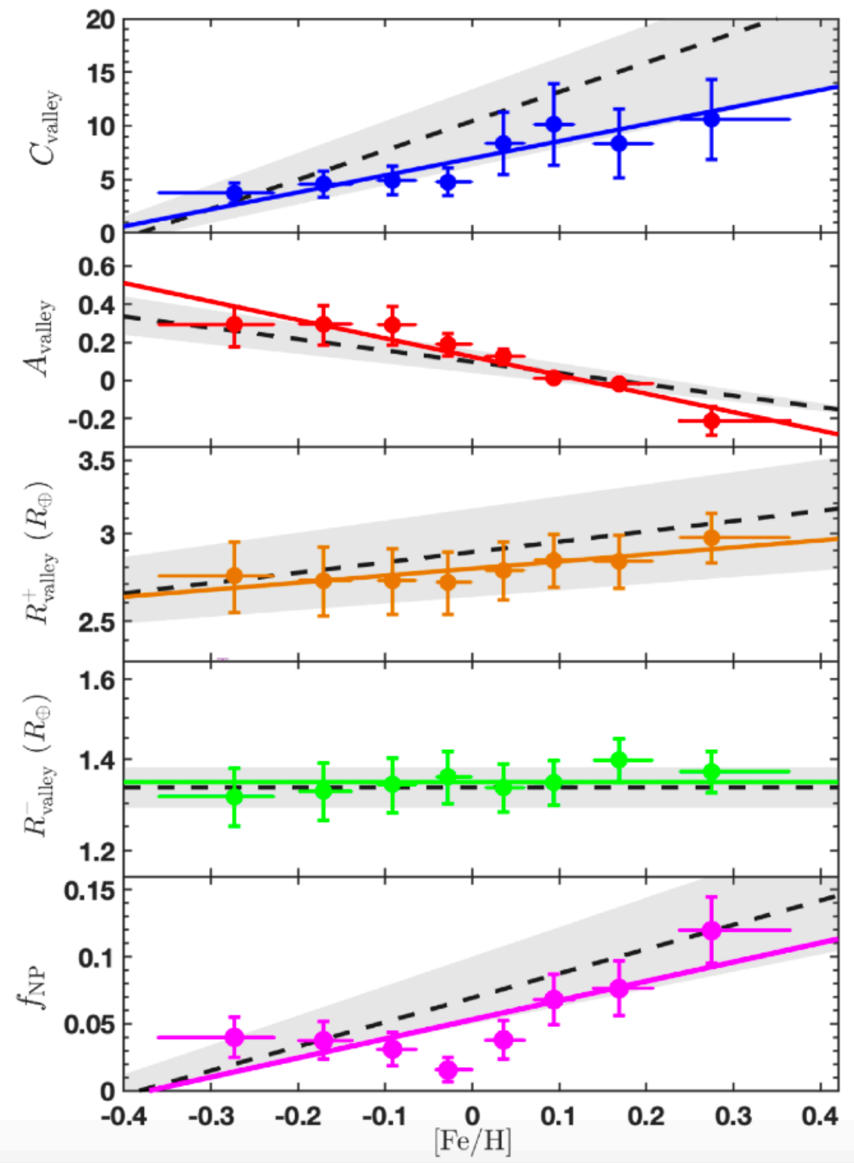
4. The average (super) Earth radius R^{-}_{valley}

$$R^{-}_{\text{valley}} = \frac{1}{N_{\text{SE}}} \sum_{i}^{N_{\text{SE}}} R_i$$

5. The Neptunian planet fraction

$$f_{\text{NP}} = \frac{N_{\text{NP}}}{N_{\text{SE}} + N_{\text{VP}} + N_{\text{SN}} + N_{\text{NP}}}$$

Dashed black lines and gray regions: observed (Chen+2022)
Coloured lines: synthetic (this work, t = 2 Gyr)

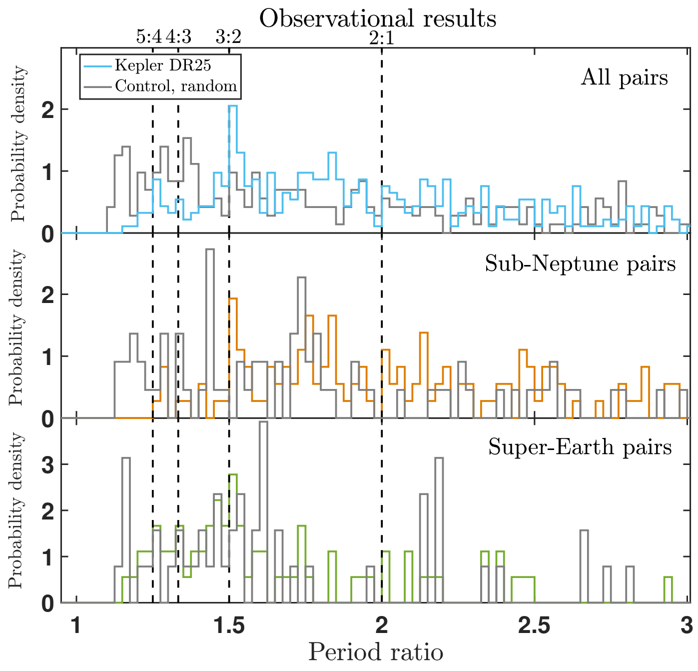


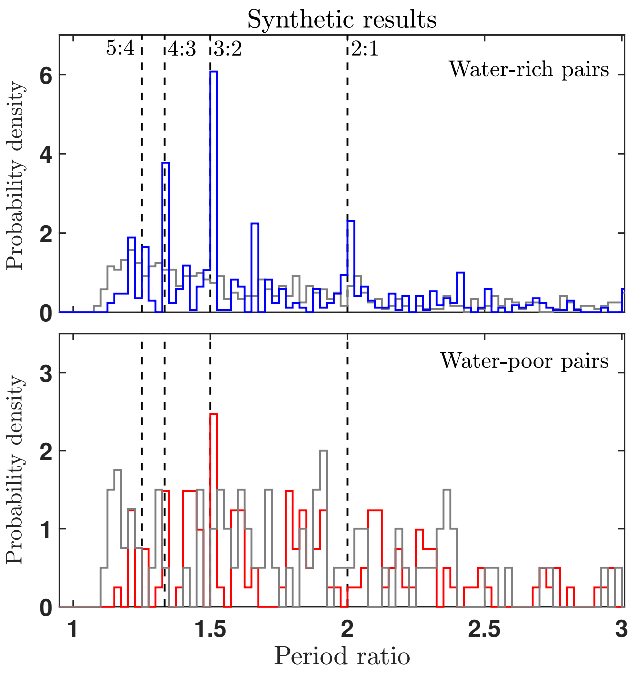
NGPPS VIII (Chen + acc.)

Conclusion: Also quantitatively good agreement - but what about the temporal evolution?

Observable imprint in period ratios ?

Is the frequency of (near) resonant pair different for super-Earth pairs different than for sub-Neptune pairs?





Chen et al. 2024

Period ratio distribution of adjacent pairs

Fraction of pairs near-MMR, normalised to random control sample

$$\hat{F}_{\mathrm{MMR}} = \frac{F^{\mathrm{obs}}_{\mathrm{MMR}}}{F^{\mathrm{con}}_{\mathrm{MMR}}}$$

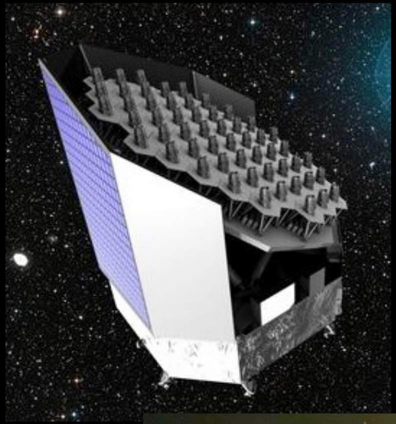
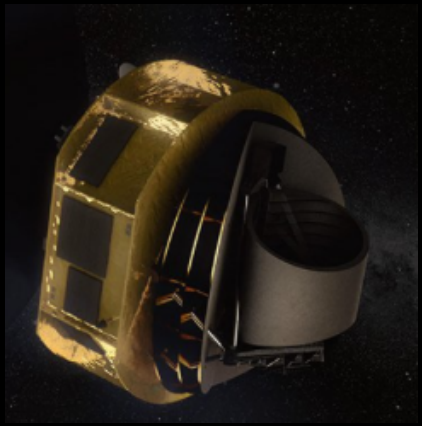
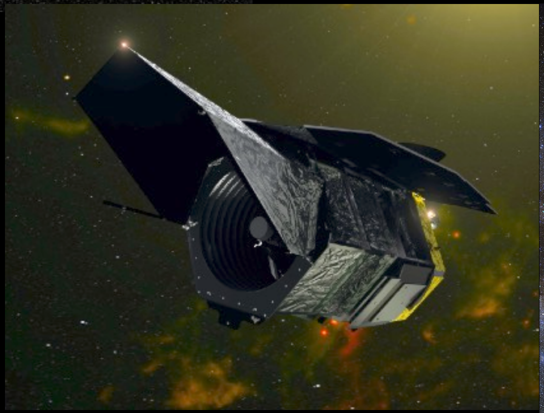
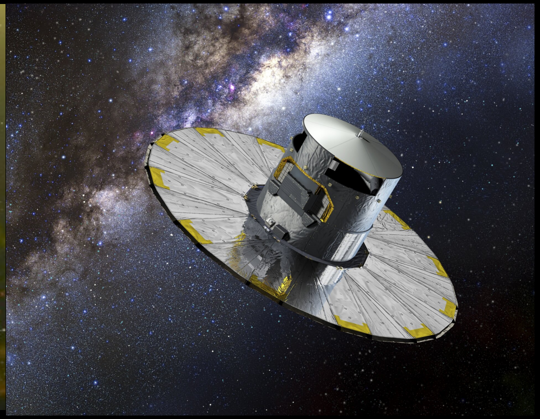
$\hat{F}_{\mathrm{MMR}}$	Kepler	Synthetic
SE-SE	$1.2^{+0.1}_{-0.2}$	$0.9^{+0.2}_{-0.2}$
SN-SN	$1.6^{+0.2}_{-0.2}$	$2.1^{+0.1}_{-0.1}$

Values >1 indicate preference for MMRs.

- SE-SE pairs: no preference (obs. and synth.)
- SN-SN pairs: some preference, stronger in synth. compared to obs.

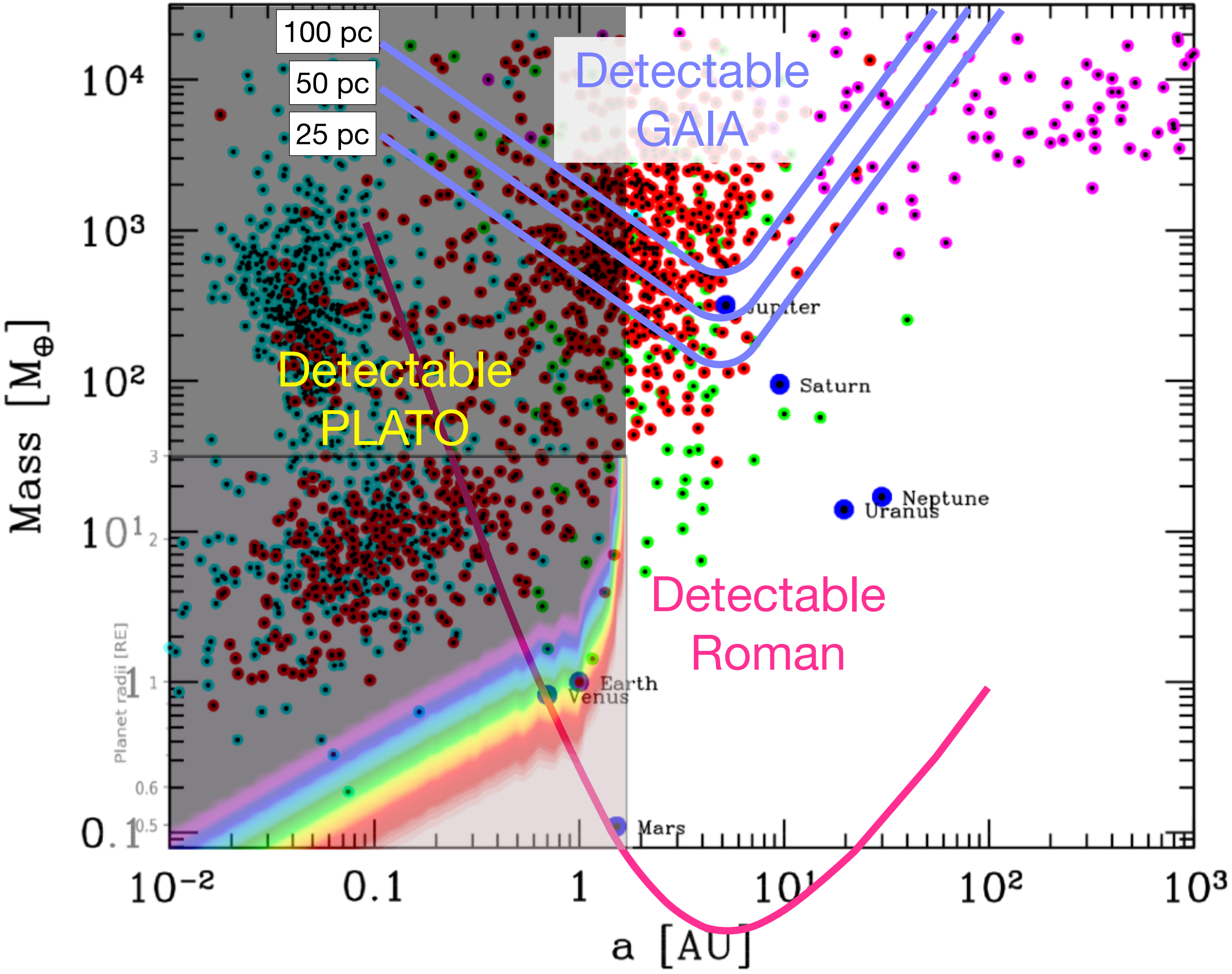
Hint that migration more important for SN than SE, but less than in model. Improve statistics with PLATO.

Alternative explanation: evolutionary effect: evaporative envelope mass loss can break MMRs (Matsumoto & Ogihara 2020; Wang & Lin 2023).

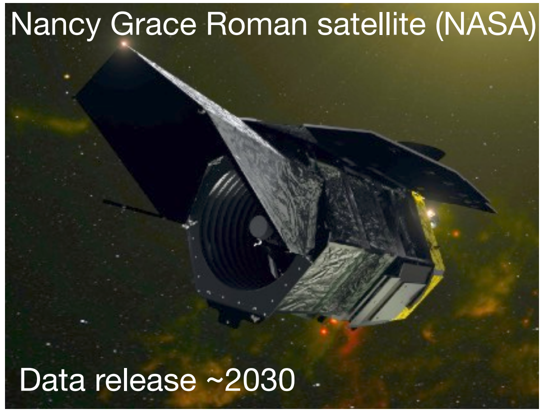
5.
Perspectives and conclusions

Outlook: future statistical exoplanet missions

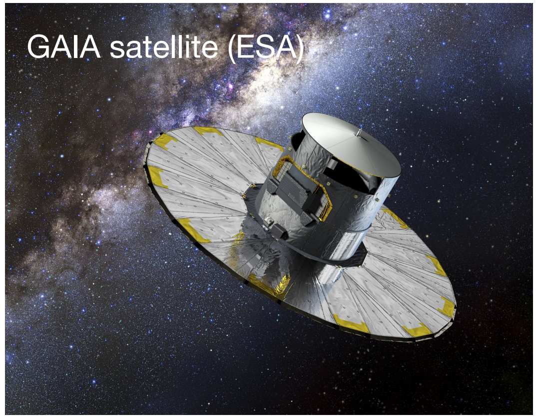


Blue lines: 5 σ detection limits for GAIA (Courtesy D. Segransan, Geneva Obs.)

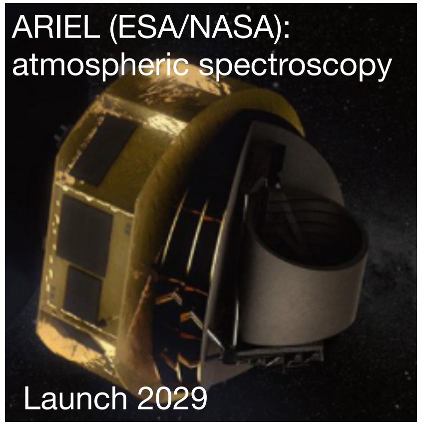
Expected yield:
4000-12000 planets



Microlensing technique
Expected yield: several thousand cold low-mass planets



Astrometric technique
Expected yield: thousands of giant exoplanets



But also new ground-based surveys

Future steps

2025

2030ies

Exoplanets and Solar System planets and are key science cases for several of the largest observational projects (space/ground) in the coming years / decades uncovering unexplored parameter spaces: key to understand if our theories of the origin of planets capture the governing physics. In this, surveys with well-defined large samples and known biases yielding the underlying demographic are of paramount importance.

Conclusions

- Population synthesis is a tool to compare exoplanet demographics and theory to improve our understanding of planet formation and evolution
  - use full wealth of observational constraints
  - put detailed models to the test
  - see global demographic consequences: which processes are key?
- Yields observational constraints for many physical mechanism
  - solid and gas accretion rate
  - N-body dynamics, tides
  - orbital migration rates
- Several observed demographical features can be reproduced, in part also quantitatively; the differences point at areas were our understanding is not complete
- Predict yield of future instruments/space missions
- Continuously improving models
  - population syntheses depend on progress of formation theory as a whole
  - many new theoretical developments to test, many new obs. constraints to come

Thank you for the attention

Some resources and further reading

Population synthesis review papers

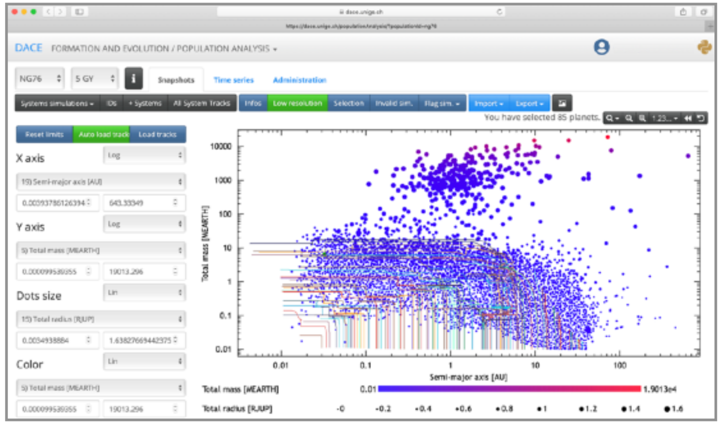
-Benz et al., Protostars & Planets VI, 691, 2014 (arXiv: 1402.7086)

-Emsenhuber et al. EPJP, 138, 2023 (arXiv: 2303.00012)

-Mordasini & Burn, RIMG, 90, 55, 2024 (arXiv: 2404.15555)

-Burn & Mordasini, Handbook of Exoplanets, 2024 (arXiv: 2410.00093)

DACE data base: online Bern population synthesis models
https://dace.unige.ch/populationAnalysis/?populationId=6



All NGPPS data publicly available via dedicated interactive online tool on DACE website

Freely available toy population synthesis model based on Ida & Lin 2004
http://nexsci.caltech.edu/workshop/2015/#handson